BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

As rescanning documents *will not* correct images, please do not report the images to the Image Problem Mailbox.

Proc. Natl. Acad. Sci. USA
Vol. 93, pp. 10614–10619, October 1996
Biochemistry

Parallel human genome analysis: Microarray-based expression monitoring of 1000 genes

(Human Genome Project/DNA chip/gene discovery/T cell)

MARK SCHENA*†, DARI SHALON‡, RENU HELLER*, ANDREW CHAI*, PATRICK O. BROWN§, AND RONALD W. DAVIS*

*Department of Biochemistry, Beckman Center, Stanford University Medical Center, Stanford, CA 94305; ‡Synteni, Palo Alto, CA 94306; and §Department of Biochemistry and Howard Hughes Medical Institute, Beckman Center, Stanford University Medical Center, Stanford, CA 94305

Contributed by Ronald W. Davis, June 26, 1996

ABSTRACT **Microarrays containing 1046 human cDNAs of unknown sequence were printed on glass with high-speed robotics. These 1.0-cm² DNA "chips" were used to quantitatively monitor differential expression of the cognate human genes using a highly sensitive two-color hybridization assay. Array elements that displayed differential expression patterns under given experimental conditions were characterized by sequencing. The identification of known and novel heat shock and phorbol ester-regulated genes in human T cells demonstrates the sensitivity of the assay. Parallel gene analysis with microarrays provides a rapid and efficient method for large-scale human gene discovery.**

Biology has entered the genome era (1). Complete genome sequences for all of the model organisms and human will probably be available by the year 2003 (2). Torrents of human expressed sequence tags (ESTs) provide a starting point for elucidating the function of tens of thousands of cognate genes (3). Genome analysis will provide insights into growth, development, differentiation, homeostasis, aging, and the onset of diseases (1–3). A detailed understanding of the human genome will require the implementation of sophisticated methods for gene expression analysis and gene discovery.

Recently, a microarray-based method for high-throughput monitoring of plant gene expression was described (4). This "chip"-based approach involved using microarrays of cDNA clones as gene-specific hybridization targets to quantitatively measure expression of the corresponding plant genes (4, 5). A two-color fluorescence labeling and detection scheme facilitated sensitive differential expression analysis of different plant tissues (4, 5). The efficiency of this approach for studies in higher plants suggested the use of this method for human genome analysis (4–7). Here, we report the use of cDNA microarrays for human gene expression monitoring, biological investigation, and gene discovery.

MATERIALS AND METHODS

Human cDNA Clones. The cDNA library was made with mRNA from human peripheral blood lymphocytes transformed with the Epstein–Barr virus. Inserts >600 bp were cloned into the lambda vector λYES-R to generate 10^7–10^8 recombinants. Bacterial transformants were obtained by infecting *E. coli* strain JM107/λKC. Colonies were picked at random and propagated in a 96-well format, and minilysate DNA was prepared by alkaline lysis using REAL preps (Qiagen, Chatsworth, CA). Inserts were amplified by PCR in a 96-well format using primers (PAN132, 5'-CCTCTATACTTTAACGTCAAGG; and PAN133, 5'-TTGTGTGGAATTGTGAGCGG) complementary to the λYES polylinker and containing a six-carbon amino modification (Glen Research, Sterling, VA) on the 5' end. PCR products were purified in a 96-well format using QIAquick columns (Qiagen).

Microarray Preparation. Amino-modified PCR products were suspended at a concentration of 0.5 mg/ml in 3× standard saline citrate (SSC) and arrayed from 96-well microtiter plates onto silylated microscope slides (CEL Associates, Houston) using high-speed robotics (4–7). A total of 1056 cDNAs, representing 1046 human clones and 10 *Arabidopsis* controls, were arrayed in 1.0-cm² areas. Printed arrays were incubated for 4 hr in a humid chamber to allow rehydration of the array elements and rinsed, once in 0.2% SDS for 1 min, twice in H_2O for 1 min, and once for 5 min in sodium borohydride solution (1.0 g of $NaBH_4$ dissolved in 300 ml of PBS and 100 ml of 100% ethanol). The arrays were submerged in H_2O for 2 min at 95°C, transferred quickly into 0.2% SDS for 1 min, rinsed twice in H_2O, air dried, and stored in the dark at 25°C.

Fluorescent Probes. Tissue mRNAs were purchased (CLONTECH). Jurkat mRNA was isolated as described by Schena *et al.* (4). Probes were made as described (4) with several modifications. The reverse transcriptase used here was Superscript II RNase H− (GIBCO). The Cy5-dCTP was purchased from Amersham. Each reverse transcription reaction contained 3.0 μg of total human mRNA. *Arabidopsis* control mRNAs were made by *in vitro* transcription of cloned HAT4, HAT22, and YesAt-23 cDNAs (4, 8, 9) using an RNA Transcription Kit (Stratagene). For quantitation, the mRNAs were doped into the reverse transcription reaction at ratios of 1:100,000, 1:10,000, and 1:1000 (wt/wt) respectively. Following the reverse transcription step, samples were treated with 2.5 μl of 1 M sodium hydroxide for 10 min at 37°C, then neutralized by adding 2.5 μl of 1 M Tris·HCl (pH 6.8) and 2.0 μl of 1 M HCl. Probe mixtures contained cDNA products derived from 3 μg of total mRNA, suspended in 5.0 μl of hybridization buffer (5× SSC plus 0.2% SDS).

Hybridization and Scanning. Probes were hybridized to 1.0-cm² microarrays under a 14 × 14 mm glass coverslip for 6–12 hr at 60°C in a custom-built hybridization chamber (4–7). Arrays were washed for 5 min at room temperature (25°C) in low stringency wash buffer (1× SSC/0.2% SDS), then for 10 min at room temperature in high stringency wash buffer (0.1× SSC/0.2% SDS). Arrays were scanned in 0.1× SSC using a fluorescence laser scanning device (4–7), fitted with a custom filter set (Chroma Technology, Brattleboro, VT). Accurate differential expression measurements (i.e., final fluorescence ratios) were obtained by taking the average of the ratios of two independent hybridizations.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviation: EST, expressed sequence tag.

Data deposition: The sequences reported in this paper have been deposited in the GenBank data base (accession nos. U56654–U56660).

†To whom reprint requests should be addressed. e-mail: schena@cmgm.stanford.edu.

Cell Culture. Jurkat cells were grown in a tissue culture incubator (37°C and 5% CO_2) in RPMI medium supplemented with 10% fetal bovine serum, 100 μg of streptomycin per ml, and 500 units of penicillin per ml. Heat shock corresponded to a 4-hr incubation at 43°C. Phorbol ester treated cells were grown for 4 hr in the presence of 50 ng of phorbol 12-myristate 13-acetate (PMA) per ml.

RNA Blotting. Dot blots were performed as described (4).

DNA Sequencing. Sequences were obtained using the PAN132 and PAN133 primers and a 373A automated sequencer, according to the instructions of the manufacturer (Applied Biosystems).

Computer Graphics and Informatics. Pseudocolor representations of fluorescent images were made with National Institutes of Health IMAGE software (version 1.52). Software for differential expression representations was purchased from Imaging Research (St. Catherine's, ON, Canada). Sequence searches were made to the nonredundant nucleotide data base at the National Center for Biotechnology Information (NCBI) using Macintosh BLAST software. The EST data base was accessed via the World Wide Web (http://www.ncbi.nlm.nih.gov/).

RESULTS

Gene Discovery and the Heat Shock Response. Microarrays were used to examine the heat shock response in cultured human T (Jurkat) cells. Control (37°C) and heat-treated (43°C) cells were harvested and lysed, and total mRNA from the two cell samples was labeled by reverse transcriptase incorporation of fluorescein- and Cy5-dCTP, respectively. In a second set of labeling reactions, the fluorescent groups were "swapped" such that samples from control and heat-treated samples were labeled with Cy5- and fluorescein-dCTP, respectively. Each pair of fluorescent probes was hybridized to a 1056-element microarray. The arrays were washed at high stringency and scanned with a confocal laser scanning device to detect emission of the two fluorescent groups.

Hybridization signals were observed to >95% of the human cDNA array elements, but not to any of the *Arabidopsis* negative controls (Fig. 1). Fluorescence intensities spanned more than three orders of magnitude for the 1046 array elements surveyed (Fig. 1). Comparative expression analysis of heat shocked versus control cells in the two experiments revealed 17 array elements that displayed altered fluorescence ratios of ≥2.0-fold (Figs. 1 and 2*A*). Of the 17 putative differentially expressed genes, 11 were induced by heat shock treatment and 6 displayed modest repression (Figs. 1 and 2*A*).

To determine the identity of the heat-regulated genes, cDNAs corresponding to each of the 17 array elements were sequenced on the proximal and distal end. Data base searches revealed perfect matches for 14 of the 17 clones, and in each case proximal and distal cDNA sequences mapped to the same gene (Table 1). Of the 1046 human genes examined on the microarray, the five most highly induced in heat-treated cells were heat shock protein 90α (hsp90α), dnaJ, hsp90β, polyubiquitin, and t-complex polypeptide-1 (tcp-1) (Table 1). Three of the 17 clones did not match any entry in the public data base, though one of the clones (B7) exhibited significant homology to an EST from *Caenorhabditis elegans* (Table 1). Each of the novel sequences (B7-B9) exhibited ≈2-fold induction (Table 1) and relatively low-level expression (Table 2).

To confirm the microarray results, mRNA levels for each of the genes were measured by RNA blotting. Each of the genes that displayed heat shock induction, including the three novel

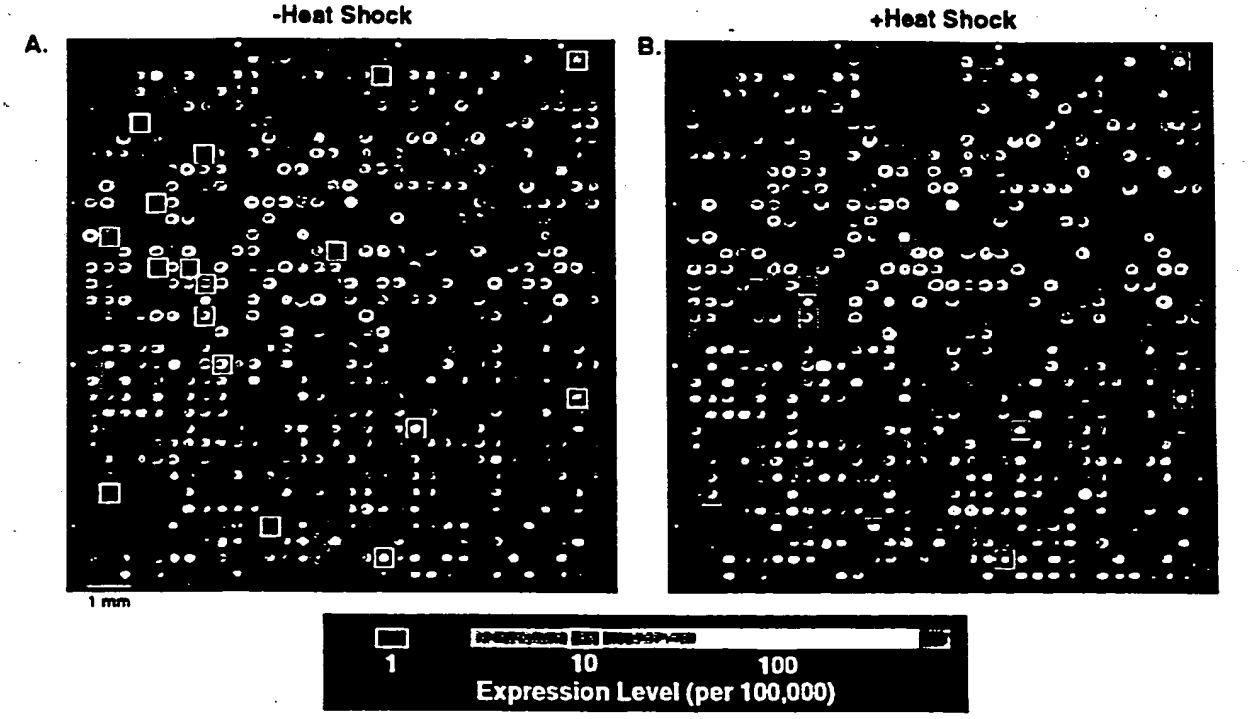


FIG. 1. Human gene expression monitored on a microarray. Fluorescent scans represented in a pseudocolor scale correspond to expression levels. The array contains 10 *Arabidopsis* controls (upper left corner, elements 1–10) and 1046 human peripheral blood cDNAs. Fluorescent probes were prepared by labeling mRNA from Jurkat cells grown at 37°C (−Heat Shock, *A*) or 43°C (+Heat Shock, *B*). Array elements that display altered fluorescence intensity (white boxes) corresponded to genes activated (red boxes) or repressed (green boxes) by heat shock. The color bar was calibrated in separate experiments using known quantities (wt/wt) of *Arabidopsis* control mRNAs added to the labeling reaction. Microarray rows (at left) and columns (at the top) are demarcated at 10 element increments (white circles). (Bar = 1 mm.)

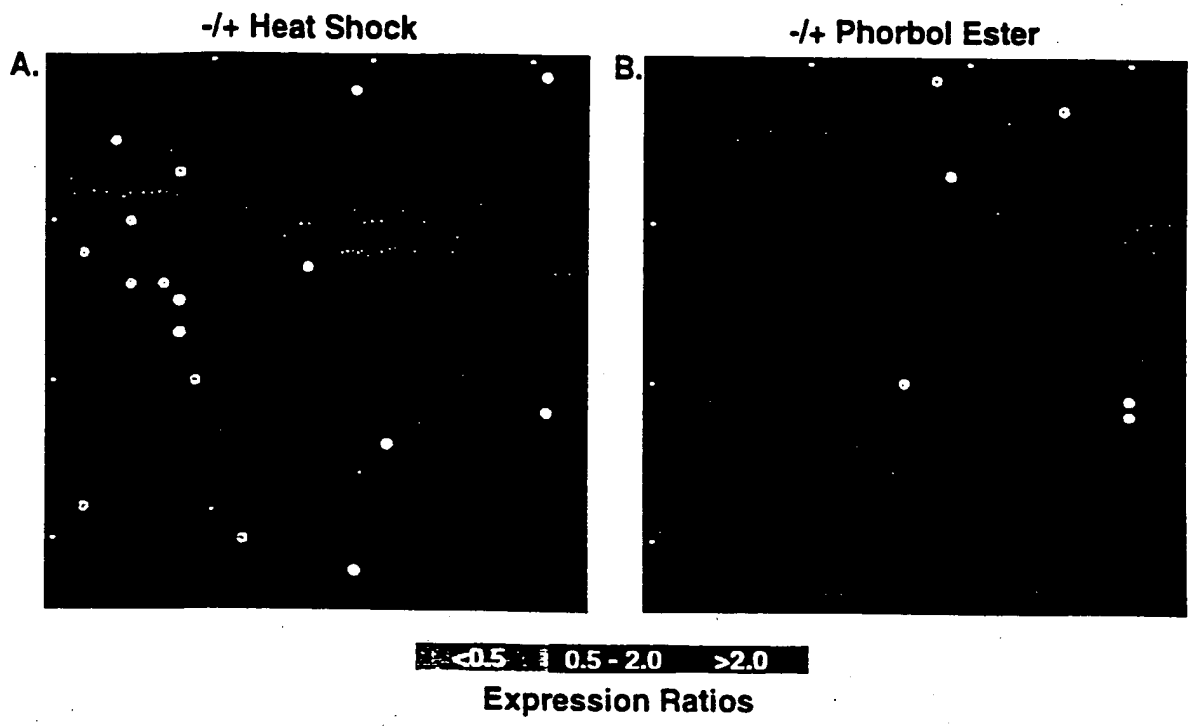


FIG. 2. Elemental displays of activated and repressed genes. Fluorescence ratios of two-color microarray scans (Fig. 1) are depicted schematically. Fluorescein-labeled probes from Jurkat cells subjected to (*A*) heat shock or (*B*) phorbol ester treatment were compared with Cy5-labeled probes from untreated cells. In a second set of reactions, the fluorescent groups were swapped (see text). The data represent the average of the ratios from two hybridizations, excluding values in which the difference of the two ratios was greater than half the average ratio. The color bar corresponds to expression ratios, which are independent of the absolute expression level of a given gene.

Table 1. Microarray elements corresponding to differentially expressed genes

Clone	Row	Column	Ratio	Blast identity	Accession no.
B1	24	21	0.5	CYC oxidase III	J01415, J01415
B2	1	31	0.5	β-Actin	NR, X00351
B3	15	8	0.5	CYC oxidase III	J01415, J01415
B4	32	19	0.5	CYC oxidase III	J01415, J01415
B5	17	8	0.5	CYC oxidase III	J01415, J01415
B6	22	31	0.5	β-Actin	NR, X00351
B7*	5	4	2.0	Novel†	U56653, U56654
B8	2	19	2.0	Novel†	U56655, U56656
B9	14	5	2.2	Novel†	U56657, U56658
B10	7	8	2.4	Polyubiquitin	X04803, X04803
B11	12	2	2.4	TCP-1	X52882, X52882
B12	28	2	2.5	Polyubiquitin	M17597, M17597
B13	14	7	2.5	Polyubiquitin	X04803, X04803
B14	20	9	2.6	HSP90β	M16660, M16660
B15	30	12	4.0	DnaJ homolog	D13388, D13388
B16	10	5	5.8	HSP90α	X07270, X07270
B17	13	16	6.3	HSP90α	M27024, X15183
B18	7	19	2.0	β_2-microglobulin	S54761, M30683
B19	21	30	2.1	Novel†	U56659, U56660
B20	3	26	2.2	β_2-microglobulin	S54761, M30683
B21	1	18	2.6	PGK	M11968, L00160
B22	22	30	3.5	NF-κB1	Z47744, M55643
B23	20	16	19	PAC-1	L11329, L11329

Clone name, array position (Fig. 1), fluorescence ratio, sequence identity, and acession number of cDNAs that manifested a differential expression pattern with probes prepared from heat shock- (B1-17) or phorbol ester-treated (B18-23) Jurkat cells. Clones showing >98% identity over 300 nucleotides were assumed to be identical to known sequences. All genes are nuclear except CYC oxidase III (mitochondrial). Accession numbers reflect the highest score for proximal and distal sequence traces, respectively. CYC, cytochrome *c*; TCP-1, T-complex polypeptide; HSP, heat shock protein; PGK, phosphoglycerate kinase; NF-κB, nuclear factor-kappaB; PAC-1, phosphatase of activated cells; and NR, trace not readable due to the presence of poly(A)+ tract.

*B7 is 67% identical to an EST from *C. elegans* (D76026).

†No match in the public data bases.

Table 2. Human gene expression monitored by microarray and RNA blot analyses

Clone	Blast identity	Expression level, per 10^5 mRNAs			
		Microarray	Ratio	RNA blot	Ratio
B1	CYC oxidase III	92/46	0.5	100/80	0.8
B2	β-Actin	240/120	0.5	270/280	1.0
B3	CYC oxidase III	36/18	0.5	ND	ND
B4	CYC oxidase III	76/38	0.5	ND	ND
B5	CYC oxidase III	62/31	0.5	ND	ND
B6	β-Actin	180/89	0.5	ND	ND
B7	Novel (weakly to D76026)	1.3/2.6	2.0	0.77/1.8	2.3
B8	Novel	2.0/4.0	2.0	1.5/3.4	2.3
B9	Novel	0.8/1.8	2.2	1.2/1.8	1.5
B10	Polyubiquitin	0.8/1.9	2.4	25/89	3.6
B11	TCP-1	2.3/5.5	2.4	7.1/27	3.8
B12	Polyubiquitin	0.8/2.0	2.5	ND	ND
B13	Polyubiquitin	1.7/4.3	2.5	ND	ND
B14	HSP90β	75/200	2.6	30/120	4.0
B15	DnaJ homolog	1.0/4.0	4.0	1.6/13	8.1
B16	HSP90α	0.6/3.5	5.8	3.2/29	9.1
B17	HSP90α	0.8/5.0	6.3	8.6/62	7.2
B18	β_2-microglobulin	1.0/2.0	2.0	5.4/15	2.8
B19	Novel	1.2/2.5	2.1	4.5/9.5	2.5
B20	β_2-microglobulin	2.7/5.9	2.2	ND	ND
B21	Phosphoglycerate kinase	2.4/6.2	2.6	4.7/9.2	2.0
B22	NF-KB1	1.7/6.0	3.5	0.65/4.7	7.2
B23	PAC-1	0.5/9.5	19	0.21/15	71

Shown are expression levels per 100,000 mRNAs (wt/wt) of genes assayed with a microarray (Fig. 1) or RNA blot. Ratios correspond to values from cells subjected to heat shock (B1-17) or phorbol ester treatment (B18-23) relative to untreated cells. Clone and gene names are given in Table 1. ND, not determined.

sequences, exhibited elevated mRNA levels by dot blot analysis (Table 2). In all cases, expression ratios as determined by the two procedures differed by <2-fold for the genes identified in the heat shock experiments (Table 2). The two assays differed more widely in terms of assessing absolute expression levels; nonetheless, absolute expression as monitored on a microarray typically correlated with RNA blots to within a factor of five (Table 2).

Phorbol Ester Signaling. To explore a signaling pathway distinct from the heat shock response, microarrays were used to examine the cellular effects of phorbol ester treatment. Jurkat cells were treated with phorbol ester, harvested, lysed, and used as a source of mRNA. Samples of mRNA from untreated or phorbol ester-stimulated cells were labeled with reverse transcriptase. The probes were mixed, hybridized to microarrays, and scanned for fluorescence emission of the two fluorescent groups. A total of six array elements displayed ≥2.0-fold elevated signals with probes from phorbol ester-treated cells relative to control samples (Fig. 2*B*).

To determine the identity of the phorbol ester-induced genes, clones corresponding to the six array elements were sequenced. Data base searches revealed perfect matches for five of the six sequences (Table 1). The two most highly induced genes were the *PAC-1* tyrosine phosphatase and nuclear factor-kappa B1 (*NF-κB1*); modest activation was observed for phosphoglycerate kinase and β_2-microglobulin (Table 1). One remaining clone (B19) did not match any entry in the public data base (Table 1). B19 displayed a 2.1-fold induction and, similar to the novel heat shock genes, a relatively low absolute expression level (Tables 1 and 2). All six of the phorbol ester-inducible genes displayed increased steady-state mRNA levels by RNA blotting (Table 2). *PAC-1* expression (Fig. 1; Table 2) defined a detection limit of ~1:500,000 for the assay.

Transcript Imaging in Human Tissues. To determine whether microarrays could be used to monitor expression in human tissues, probes were prepared from human bone marrow, brain, prostate, and heart by labeling each mRNA sample with Cy5-dCTP. In a separate reaction, a control probe was prepared by labeling Jurkat mRNA with fluorescein-dCTP. The four Cy5-labeled probes were each mixed with an aliquot of the fluorescein-labeled control sample, and the four mixtures were hybridized to separate microarrays. The arrays were washed and scanned for fluorescence emission, and hybridization signals for each of the tissues samples were normalized to the Jurkat control to generate an expression profile for each of the 1046 clones present on the array.

Detectable expression was observed for all 15 of the heat shock and phorbol ester-regulated genes in the four tissue types examined (Fig. 3). In general, the expression level of each gene in Jurkat cells correlated rather closely with expression in the four tissues (Table 2; Fig. 3). Genes encoding β-actin and cytochrome *c* oxidase, the two most highly expressed of the 15 genes in Jurkat cells (Table 2), were highly expressed in bone marrow, brain, prostate, and heart (Fig. 3*A*). Expression of cytochrome *c* oxidase, hsp90α, and the novel B7 sequence was significantly greater in heart than in the other tissues (Fig. 3).

DISCUSSION

Many of the heat shock genes identified in this study encode factors that function either as molecular "chaperones" (HSP90α, HSP90β, DnaJ, TCP-1) or as mediators of protein degradation (polyubiquitin). The identification of these sequences is consistent with the biochemical basis of heat shock induction (10–15). Proteins undergo denaturation at elevated temperatures, and those that fail to maintain proper conformation must be selectively degraded (10–15). It will be interesting to determine whether the three novel heat shock-inducible sequences (B7–B9) mediate protein folding and turnover or possess some other biochemical activity. Complete nucleotide sequence determination, conceptual translation, expression monitoring, and biochemical analysis should provide a detailed functional understanding of these genes.

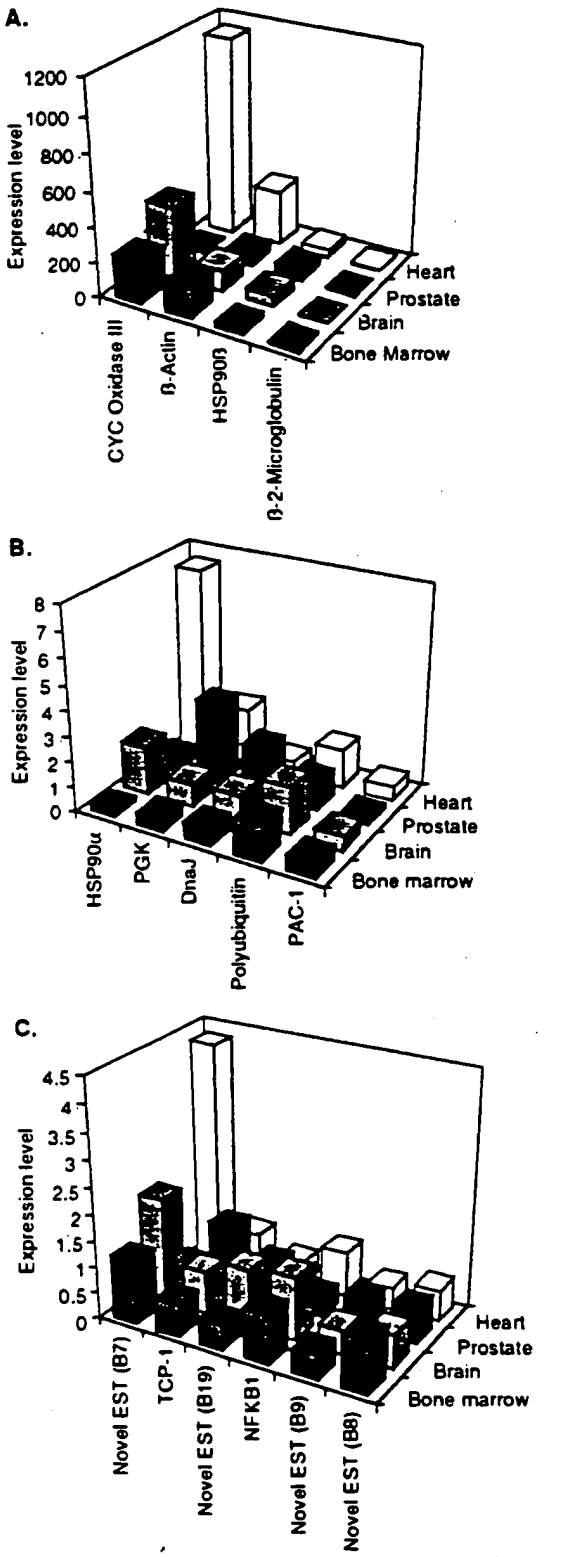


FIG. 3. Transcript profiles of heat shock and phorbol ester-regulated genes. Gene expression levels per 100,000 mRNAs (*x*-axes) are shown for 15 genes (Table 1) in human bone marrow (red), brain (green), prostate (blue), and heart (yellow). Genes are grouped according to expression levels (*A–C*).

Phorbol ester, a potent activator of protein kinase C (16, 17), induced a set of genes distinct from those involved in the heat shock pathway. The most highly induced gene identified in this study, *PAC-1*, encodes a nuclear tyrosine kinase that may play a role in regulating transcription and cell cycle progression (18). NF-κB1, a second phorbol ester-inducible gene, is an intensively studied member of the Rel transcription factor family (19–21). The Rel proteins are activated by a large number of stimuli, including phorbol esters, cytokines, bacterial and viral pathogens, and ultraviolet light (19–21). Modest activation was observed for three sequences not known to be inducible by phorbol esters, including phosphoglycerate kinase, β_2-microglobulin, and a novel human gene (B19). Extensive expression monitoring with microarrays should assist in understanding how each of these genes integrate into the highly complex phorbol ester signaling pathway.

It is striking that four novel human genes were discovered with an array of 1000 randomly chosen clones, particularly because the heat shock and phorbol ester signaling pathways have been so intensively studied (10–21). The facile discovery of these sequences underscores the fact that microarrays can be used for gene discovery in the absence of any sequence information. By this approach, clones are chosen at random from any library of interest and only those clones that display interesting expression patterns are sequenced and characterized. This parallel assay, coupled with a modest DNA sequencing facility, allows high-throughput human genome expression analysis and gene discovery.

Genes that are activated or repressed by a given stimulus provide functional clues to the cellular pathway involved (22–24). Detailed examination of these gene expression "signatures" can provide a dynamic view of the mode of action of a given signaling substance (22–24). Microarrays may thus allow rapid mechanistic examination of hormones, drugs, elicitors, and other small molecules; moreover, functional analysis of transcription factors, kinases, growth factors, cytokines, receptors, and other gene products should be possible. Efforts are underway to develop mRNA amplification strategies to enable probe preparation from minute tissue samples. This capability might allow for high-throughput patient screening in a clinical setting.

The current detection limit of the assay allows monitoring of transcripts that represent ≈1:500,000 (wt/wt) of the total mRNA. This 10-fold increase in sensitivity compared with the original report (4) was achieved largely by modifying the coupling chemistry, which reduced background fluorescence. The significance of this improvement is considerable in that approximately half the human genes identified in this study, including all four novel sequences, exhibited expression levels below the original detection limit of 1:50,000 (4).

The ability to detect 2-fold changes in expression was achieved by the use of two-color fluorescence in the labeling and detection schemes, digitized data collection, and custom software. The importance of this capability is underscored by the fact that nearly all of the genes examined here exhibited <6-fold changes in expression. The four novel genes, which showed ≤2.2-fold activation, were probably overlooked in previous screens that used conventional differential expression techniques. It may be possible to further improve the precision of the microarray assay by the use of closely related fluorescent analogs, such as Cy3 and Cy5, in the labeling and hybridization reactions.

Microarrays offer a number of advantages over other potential high-capacity approaches to expression analysis. The chip-based approach enables small hybridization volumes, high array densities, and the use of fluorescence labeling and detection schemes. These features provide a set of perf r-mance specifications that are unattainable with filter-based approaches (25, 26). The use of cDNA clones provides hybridization specificity that is not readily attained with oligo-

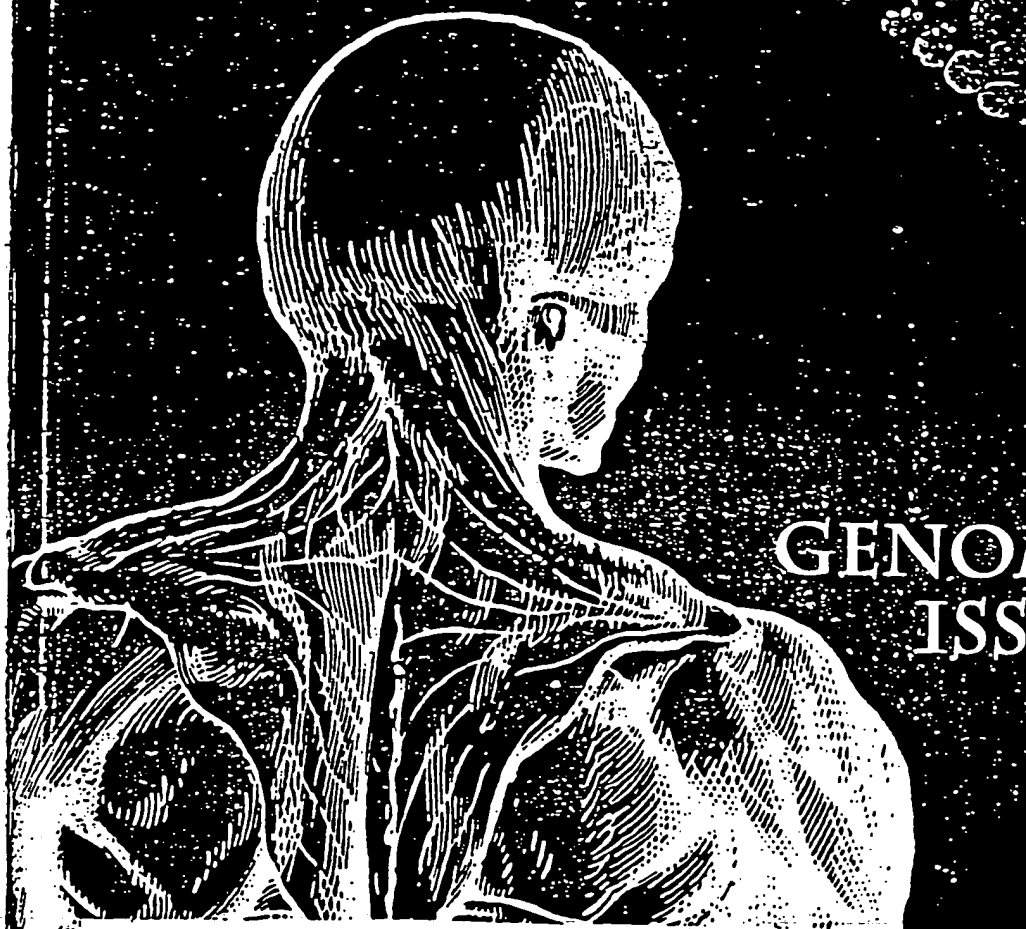
nucleotide arrays (27–30). The parallel format of the assay provides a simultaneous differential expression readout for >1000 genes. This contrasts with sequencing-based methods, which require serial data collection for expression analysis (31, 32). A commercial source of cDNA microarrays would greatly speed the use of a chip-based approach to expression analysis.

The availability of large numbers of ESTs (3) provides a rich resource of human cDNA clones for microarraying. The >400,000 ESTs in the public data bases represent a significant subset of all human genes (3, 33). Microarrays of thousands of ESTs will provide a powerful analytical tool for future human gene expression studies. The ≈100,000 genes in the human genome (2, 33) emphasize the need for microarrays of greater density. Attempts to improve microdeposition techniques are underway and should allow construction of arrays containing a complete set of human gene targets (http://cmgm.stanford.edu/~schena/). Microarrays of ≈100,000 cDNA elements would allow expression monitoring of the entire human genome in a single hybridization. This capacity, coupled with detailed biochemical analysis of the individual gene products, would greatly speed the functional analysis of the human genome.

We thank S. Elledge (selledge@bcm.tmc.edu) for the human cDNA library, Qiagen representatives for help with plasmid purification, and A. J. Smith and colleagues at the Protein and Nucleic Acid (PAN) facility (Stanford) for oligonucleotide synthesis and DNA sequencing. We also thank members of the Davis, Brown, and Smith laboratories for critical comments and helpful discussions and Synteni employees for technical assistance. Support for R.W.D. was provided by the National Science Foundation (MCB9106011) and National Institutes of Health (R37HG00198) and for P.O.B. by the National Institutes of Health (3R21HG00450) and Howard Hughes Medical Institute. P.O.B. is an assistant investigator of the Howard Hughes Medical Institute.

1. Watson, J. D. (1993) *Gene* **135**, 309–315.
2. Collins, F. S. (1995) *Proc. Natl. Acad. Sci. USA* **92**, 10821–10823.
3. Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie, W. R. & Venter, J. C. (1991) *Science* **252**, 1651–1656.
4. Schena, M., Shalon, D., Davis, R. W. & Brown, P. O. (1995) *Science* **270**, 467–470.
5. Shalon, D. (1996) Ph.D. thesis (Stanford University).
6. Schena, M. (1996) *BioEssays* **18**, 427–431.
7. Shalon, D., Smith, S. J. & Brown, P. O. (1996) *Genome Res.* **6**, 639–645.
8. Schena, M. & Davis, R. W. (1994) *Proc. Natl. Acad. Sci. USA* **91**, 8393–8397.
9. Schena, M. & Davis, R. W. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 3894–3898.
10. Jindal, S. (1996) *Trends Biotechnol.* **14**, 17–20.
11. Wilkinson, K. D. (1995) *Annu. Rev. Nutr.* **15**, 161–189.
12. Jakob, U. & Buchner, J. (1994) *Trends Biochem. Sci.* **19**, 205–211.
13. Becker, J. & Craig, E. A. (1994) *Eur. J. Biochem.* **219**, 11–23.
14. Cyr, D. M., Langer, T. & Douglas, M. G. (1994) *Trends Biochem. Sci.* **19**, 176–181.
15. Craig, E. A., Weissman, J. S. & Horwich, A. L. (1994) *Cell* **78**, 365–372.
16. Newton, A. C. (1995) *J. Biol. Chem.* **270**, 28495–28498.
17. Nishizuka, Y. (1995) *FASEB J.* **9**, 484–496.
18. Rohan, P. J., Davis, P., Moskaluk, C. A., Kearns, M., Krutzsch, H., Siebenlist, U. & Kelly, K. (1993) *Science* **259**, 1763–1766.
19. Thanos, D. & Maniatis, T. (1995) *Cell* **80**, 529–532.
20. Baeuerle, P. A. & Henkel, T. (1994) *Annu. Rev. Immunol.* **12**, 141–179.
21. Liou, H.-C. & Baltimore, D. (1993) *Curr. Opin. Cell Biol.* **5**, 477–487.
22. Cohen, G. B., Ren, R. & Baltimore, D. (1995) *Cell* **80**, 237–248.
23. Chan, A. C., Desai, D. M. & Weiss, A. (1994) *Annu. Rev. Immunol.* **12**, 555–592.
24. Crabtree, G. R. & Clipstone, N. A. (1994) *Annu. Rev. Biochem.* **63**, 1045–1083.
25. Gress, T. M., Hoheisel, J. D., Lennon, G. G., Zehetner, G. & Lehrach, H. (1992) *Mamm. Genome* **3**, 609–619.
26. Bernard, K., Auphan, N., Granjeaud, S., Victorero, G., Schmitt-Verhulst, A.-M., Jordan, B. R. & Nguyen, C. (1996) *Nucleic Acids Res.* **24**, 1435–1442.
27. Fodor, S. P. A., Read, J. L., Pirrung, M. C., Stryer, L., Lu, A. T. & Solas, D. (1991) *Science* **251**, 767–773.
28. Southern, E. M., Maskos, U. & Elder, J. K. (1992) *Genomics* **13**, 1008–1017.
29. Guo, Z., Guilfoyle, R. A., Thiel, A. J., Wang, R. & Smith, L. M. (1994) *Nucleic Acids Res.* **22**, 5456–5465.
30. Matson, R. S., Rampal, J., Pentoney, S. L., Jr., Anderson, P. D. & Coassin, P. (1995) *Anal. Biochem.* **224**, 110–116.
31. Velculescu, V. E., Zhang, L., Vogelstein, B. & Kinzler, K. W. (1995) *Science* **270**, 484–487.
32. Adams, M. D. (1996) *BioEssays* **18**, 261–262.
33. Fields, C., Adams, M. D., White, O. & Venter, J. C. (1994) *Nat. Genet.* **7**, 345–346.

SCIENCE
30 OCTOBER 1998
Vol. 282 • Pages 649–53[illegible]

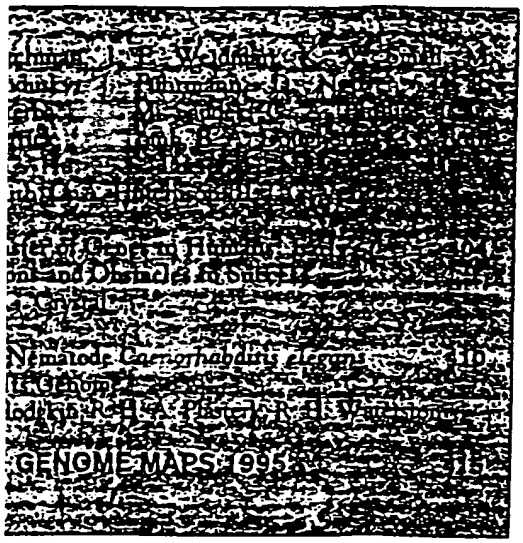


GENOME ISSUE

COVER

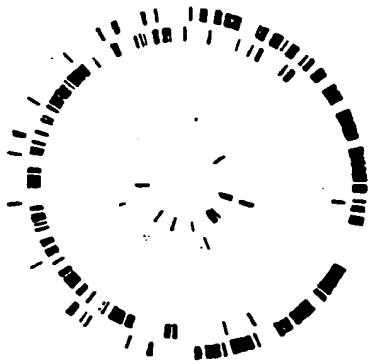
e Genome Project adds a new dimension to questions . gene expression in humans and model systems. A art on page 415 summarizes progr ss in the *enorhabditis elegans* Genome Project and indicates me ways information about sequences can be used.

News stories, Articles, Perspectives, Policy Forums, and Reports focus on technological developments, clinical applications, and ethical concerns resulting from the burgeoning of genomic information. [*C. elegans* image: F. Maduro and D. Pilgrim, University of Alberta]



EPORTS

TECHNICAL COMMENTS



397

Good things in small genomes

SCIENCE (ISSN 0036-8075) is published weekly on Friday, except the last week in December, by the American Association for the Advancement of Science, 1333 H Street, NW, Washington, DC 20005. Second-class postage (publication No. 484460) paid at Washington, DC, and additional mailing offices. Copyright © 1995 by the American Association for the Advancement of Science. The title SCIENCE is a registered trademark of the AAAS. Domestic individual membership and subscripson (51 issues): $97 ($50 allocated to subscripson). Domestic institutional subscripson (51 issues): $228. Foreign postage extra: Mexico, Caribbean (surface mail) $53; other countries (air assist delivery) $93. First class, airmail, student and emeritus rates on request. Canadian rates with GST available upon request, GST #1254 88122. Printed in the U.S.A.

Change of address: allow 4 weeks, giving old and new addresses and 8-digit account number. Postmaster: Send change of address to *Science*, P.O. Box 1811, Danbury, CT 06813-1811. Single copy sales: $7.00 per issue prepaid includes surface postage; bulk rates on request. Authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act is granted by AAAS to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that $3.00 per article is paid directly to CCC, 27 Congress Street, Salem, MA 01970. The identification code for *Science* is 0036-8075/93 $3.00. *Science* is indexed in the *Reader's Guide to Periodical Literature* and in several specialized indexes.

Indicates accompanying feature

Axl1p sequence following Ser^{206} and occurs within the domain of Axl1p that shows homology with hDE (*14*). To delete the complete *STE23* sequence and create the *ste23Δ::URA3* mutation, polymerase chain reaction (PCR) primers (5′-TCGGAAGACCTCATTCTTGCTCATTTTGATATTGCTC- TGTAGATTGTACTGAGAGTGCAC-3′; and 5′-GCTACAAACAGCGTCGACTTGAATGCCCCGACATCTTCGACTGTGCGGTATTTCACACCG-3′) were used to amplify the *URA3* sequence of pRS316, and the reaction product was transformed into yeast for one-step gene replacement [R. Rothstein, *Methods Enzymol.* **194**, 281 (1991)]. To create the *axl1Δ::LEU2* mutation contained on p114, a 5.0-kb Sal I fragment from *pAXL1* was cloned into pUC19, and an internal 4.0-kb Hpa I–Xho I fragment was replaced with a *LEU2* fragment. To construct the *ste23Δ::LEU2* allele (a deletion corresponding to 931 amino acids) carried on p153, a *LEU2* fragment was used to replace the 2.8-kb Pml I–Ecl136 II fragment of *STE23*, which occurs within a 6.2-kb Hind III–Bgl II genomic fragment carried on pSP72 (Promega). To create YEp*MFA1*, a 1.6-kb Bam HI fragment containing *MFA1*, from pKK16 [K. Kuchler, R. E. Sterne, J. Thorner, *EMBO J.* **8**, 3973 (1989)], was ligated into the Bam HI site of YEp351 [J. E. Hill, A. M. Myers, T. J. Koerner, A. Tzagoloff, *Yeast* **2**, 163 (1986)].

4. J. Chant and I. Herskowitz, *Cell* **65**, 1203 (1991).
5. B. W. Matthews, *Acc. Chem. Res.* **21**, 333 (1988).
6. K. Kuchler, H. G. Dohlman, J. Thorner, *J. Cell Biol.* **120**, 1203 (1993); R. Kolling and C. P. Hollenberg, *EMBO J.* **13**, 3261 (1994); C. Berkower, D. Loayza, S. Michaelis, *Mol. Biol. Cell* **5**, 1185 (1994).
7. A. Bender and J. R. Pringle, *Proc. Natl. Acad. Sci. U.S.A* **86**, 9976 (1989); J. Chant, K. Corrado, J. R. Pringle, I. Herskowitz, *Cell* **65**, 1213 (1991); S. Powers, E. Gonzales, T. Christensen, J. Cubert, D. Broek, *ibid.*, p. 1225; H. O. Park, J. Chant, I. Herskowitz, *Nature* **365**, 269 (1993); J. Chant, *Trends Genet.* **10**, 328 (1994); ——— and J. R. Pringle, *J. Cell Biol.* **129**, 751 (1995); J. Chant, M. Mischke, E. Mitchell, I. Herskowitz, J. R. Pringle, *ibid.*, p. 767.
8. G. F. Sprague Jr., *Methods. Enzymol.* **194**, 77 (1991).
9. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
10. A W303 1A derivative, SY2625 (*MATa ura3-1 leu2-3, 112 trp1-1 ade2-1 can1-100 sst1Δ mfa2Δ::FUS1-lacZ his3Δ::FUS1-HIS3*), was the parent strain for the mutant search. SY2625 derivatives for the mating assays, secreted pheromone assays, and the pulse-chase experiments included the following strains: Y49 (*ste22-1*), Y115 (*mfa1Δ::LEU2*), Y142 (*axl1::URA3*), Y173 (*axl1Δ::LEU2*), Y220 (*axl1::URA3 ste23Δ::URA3*), Y221 (*ste23Δ::URA3*), Y231 (*axl1Δ::LEU2 ste23Δ::LEU2*), and Y233 (*ste23Δ::LEU2*). *MATα* derivatives of SY2625 included the following strains: Y199 (SY2625 made *MATα*), Y278 (*ste22-1*), Y195 (*mfa1Δ::LEU2*), Y196 (*axl1Δ::LEU2*), and Y197 (*axl1::URA3*). The EG123 (*MATa leu2 ura3 trp1 can1 his4*) genetic background was used to create a set of strains for analysis of bud site selection. EG123 derivatives included the following strains: Y175 (*axl1Δ::LEU2*), Y223 (*axl1::URA3*), Y234 (*ste23Δ::LEU2*), and Y272 (*axl1Δ::LEU2 ste23Δ::LEU2*). *MATα* derivatives of EG123 included the following strains: Y214 (EG123 made *MATα*) and Y293 (*axl1Δ::LEU2*). All strains were generated by means of standard genetic or molecular methods involving the appropriate constructs (*23*). In particular, the *axl1 ste23* double mutant strains were created by crossing of the appropriate *MATa ste23* and *MATα axl1* mutants, followed by sporulation of the resultant diploid and isolation of the double mutant from nonparental di-type tetrads. Gene disruptions were confirmed with either PCR or Southern (DNA) analysis.
11. p129 is a YEp352 [J. E. Hill, A. M. Myers, T. J. Koerner, A. Tzagoloff, *Yeast* **2**, 163 (1986)] plasmid containing a 5.5-kb Sal I fragment of *pAXL1*. p151 was derived from p129 by insertion of a linker at the Bgl II site within *AXL1*, which led to an in-frame insertion of the hemagglutinin (HA) epitope (DOYPYDVPDYA) (*29*) between amino acids 854 and 855 of the *AXL1* product. pC225 is a KS+ (Stratagene) plasmid containing a 0.5-kb Bam HI–Sst I fragment from *pAXL1*. Substitution mutations of the proposed active site of Axl1p were created with the use of pC225 and site-specific mutagenesis involving appropriate synthetic oligonucleotides (*axl1-H68A*, 5′-GTGCTCACAAAGCGCTGCCAAACCGGC-3′; *axl1-E71A*, 5′-AAGAATCATGTGCGCACAAAGGTGCGC-3′; and *axl1-E71D*, 5′-AAGAATCATGTGATCACAAAGGTGCGC-3′). The mutations were confirmed by sequence analysis. After mutagenesis, the 0.4-kb Bam HI–Msc I fragment from the mutagenized pC225 plasmids was transferred into *pAXL1* to create a set of pRS316 plasmids carrying different *AXL1* alleles, p124 (*axl1-H68A*), p130 (*axl1-E71A*), and p132 (*axl1-E71D*). Similarly, a set of HA-tagged alleles carried on YEp352 were created after replacement of the p151 Bam HI–Msc I fragment, to generate p161 (*axl1-E71A*), p162 (*axl1-H68A*), and p163 (*axl1-E71D*).

32. We thank J. Becker and S. Michaelis for providing a-factor antibodies; S. Michaelis for discussing unpublished results and helping with the pulse-chase experiments; J. Brown, J. Chant, and S. Sanders for their input concerning bud site selection experiments; M. Raymond, F. Tamanoi, and M. Whiteway for plasmids; M. Marra for providing the *STE23* genomic fragment; and H. Bussey, J. Brown, N. Davis, T. Favero, C. de Hoog, and S. Kim for comments on the manuscript. Supported by a grant to C.B. from the Natural Sciences and Engineering Research Council of Canada. Support for M.N.A. was from a California Tobacco-Related Disease Research Program postdoctoral fellowship (4FT-0083).

22 June 1995; accepted 21 August 1995

Quantitative Monitoring of Gene Expression Patterns with a Complementary DNA Microarray

Mark Schena,*† Dari Shalon,*† Ronald W. Davis, Patrick O. Brown‡

A high-capacity system was developed to monitor the expression of many genes in parallel. Microarrays prepared by high-speed robotic printing of complementary DNAs on glass were used for quantitative expression measurements of the corresponding genes. Because of the small format and high density of the arrays, hybridization volumes of 2 microliters could be used that enabled detection of rare transcripts in probe mixtures derived from 2 micrograms of total cellular messenger RNA. Differential expression measurements of 45 *Arabidopsis* genes were made by means of simultaneous, two-color fluorescence hybridization.

The temporal, developmental, topographical, histological, and physiological patterns in which a gene is expressed provide clues to its biological role. The large and expanding database of complementary DNA (cDNA) sequences from many organisms (*1*) presents the opportunity of defining these patterns at the level of the whole genome.

For these studies, we used the small flowering plant *Arabidopsis thaliana* as a model organism. *Arabidopsis* possesses many advantages for gene expression analysis, including the fact that it has the smallest genome of any higher eukaryote examined to date (*2*). Forty-five cloned *Arabidopsis* cDNAs (Table 1), including 14 complete sequences and 31 expressed sequence tags (ESTs), were used as gene-specific targets. We obtained the ESTs by selecting cDNA clones at random from an *Arabidopsis* cDNA library. Sequence analysis revealed that 28 of the 31 ESTs matched sequences in the database (Table 1). Three additional cDNAs from other organisms served as controls in the experiments.

The 48 cDNAs, averaging ~1.0 kb, were amplified with the polymerase chain reaction (PCR) and deposited into individual wells of a 96-well microtiter plate. Each sample was duplicated in two adjacent wells to allow the reproducibility of the arraying and hybridization process to be tested. Samples from the microtiter plate were printed onto glass microscope slides in an area measuring 3.5 mm by 5.5 mm with the use of a high-speed arraying machine (*3*). The arrays were processed by chemical and heat treatment to attach the DNA sequences to the glass surface and denature them (*3*). Three arrays, printed in a single lot, were used for the experiments here. A single microtiter plate of PCR products provides sufficient material to print at least 500 arrays.

Fluorescent probes were prepared from total *Arabidopsis* mRNA (*4*) by a single round of reverse transcription (*5*). The *Arabidopsis* mRNA was supplemented with human acetylcholine receptor (AChR) mRNA at a dilution of 1:10,000 (w/w) before cDNA synthesis, to provide an internal standard for calibration (*5*). The resulting fluorescently labeled cDNA mixture was hybridized to an array at high stringency (*6*) and scanned

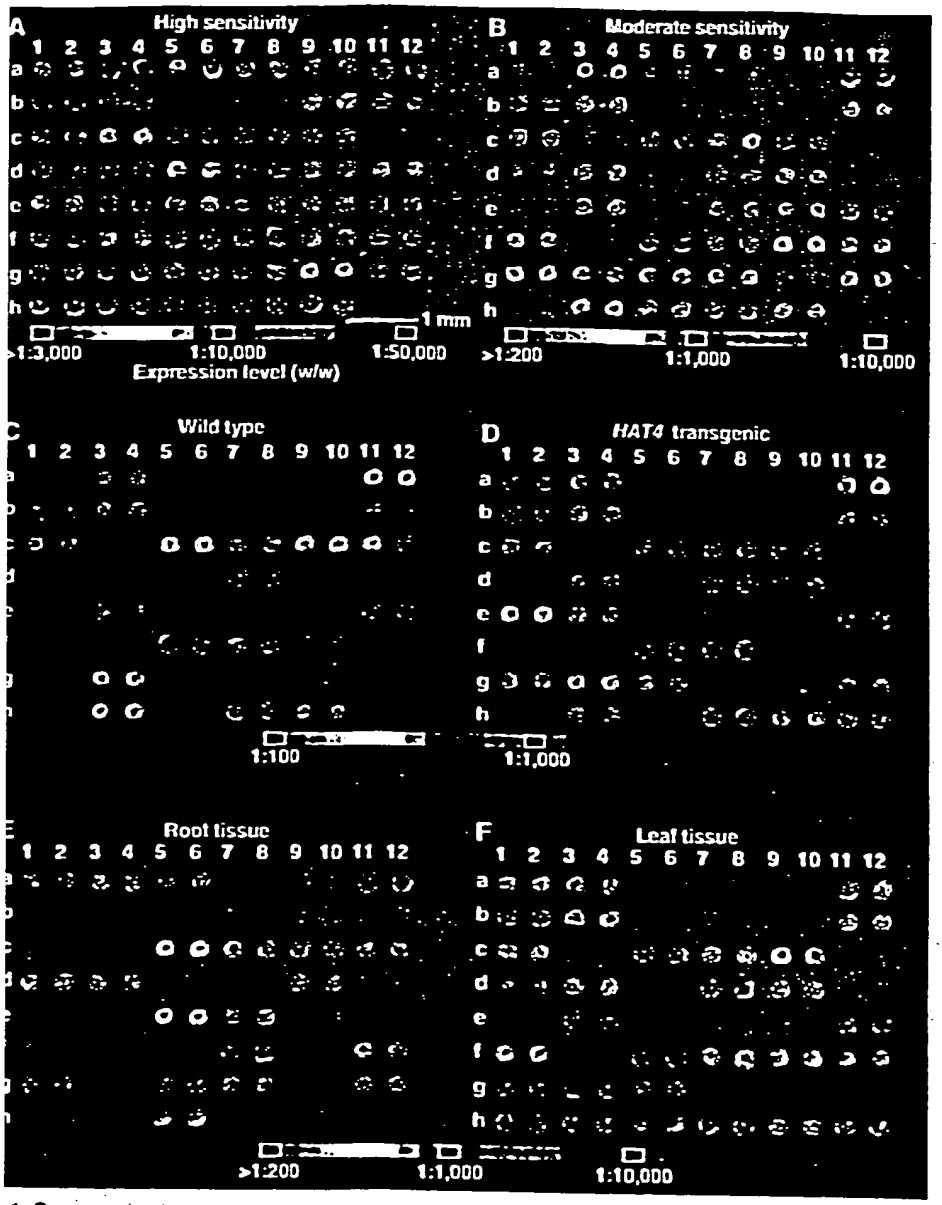
M. Schena and R. W. Davis, Department of Biochemistry, Beckman Center, Stanford University Medical Center, Stanford, CA 94305, USA.
D. Shalon and P. O. Brown, Department of Biochemistry and Howard Hughes Medical Institute, Beckman Center, Stanford University Medical Center, Stanford, CA 94305, USA.

*These authors contributed equally to this work.
†Present address: Synteni, Palo Alto, CA 94303, USA.
‡To whom correspondence should be addressed. E-mail: pbrown@cmgm.stanford.edu

:h a laser (3). A high-sensitivity scan gave nals that saturated the detector at nearly of the *Arabidopsis* target sites (Fig. 1A). ilibration relative to the AChR mRNA ndard (Fig. 1A) established a sensitivity iit of ~1:50,000. No detectable hybridiza- n was observed to either the rat glucocor- id receptor (Fig. 1A) or the yeast *TRP4* g. 1A) targets even at the highest scan- ig sensitivity. A moderate-sensitivity scan of the same array allowed linear detection of the more abundant transcripts (Fig. 1B). Quantitation of both scans revealed a range of expression levels spanning three orders of magnitude for the 45 genes tested (Table 2). RNA blots (7) for several genes (Fig. 2) corroborated the expression levels measured with the microarray to within a factor of 5 (Table 2).

Differential gene expression was investigated with a simultaneous, two-color hybridization scheme, which served to minimize experimental variation inherent in the comparison of independent hybridizations. Fluorescent probes were prepared from two mRNA sources with the use of reverse transcriptase in the presence of fluorescein- and lissamine-labeled nucleotide analogs, respectively (5). The two probes were then mixed together in equal proportions, hybridized to a single array, and scanned separately for fluorescein and lissamine emission after independent excitation of the two fluorophores (3).

To test whether overexpression of a single gene could be detected in a pool of total *Arabidopsis* mRNA, we used a microarray to analyze a transgenic line overexpressing the single transcription factor *HAT4* (8). Fluorescent probes representing mRNA from wild-type and *HAT4*-transgenic plants were labeled with fluorescein and lissamine, respectively; the two probes were then mixed and hybridized to a single array. An intense hybridization signal was observed at the position of the *HAT4* cDNA in the lissamine-specific scan (Fig. 1D), but not in the fluorescein-specific scan of the same array (Fig. 1C). Calibration with AChR mRNA added to the fluorescein and lissamine cDNA synthesis reactions at dilutions of 1:10,000 (Fig. 1C) and 1:100 (Fig. 1D), respectively, revealed a 50-fold elevation of *HAT4* mRNA in the transgenic line relative to its abundance in wild-type plants (Table 2). This magnitude of *HAT4* overexpression matched that inferred from the Northern (RNA) analysis within a factor of 2 (Fig. 2 and Table 2). Expression of all the other genes monitored on the array differed by less than a factor of 5 between *HAT4*-transgenic and wild-type plants (Fig 1, C



1. Gene expression monitored with the use of cDNA microarrays. Fluorescent scans represented in udocolor correspond to hybridization intensities. Color bars were calibrated from the signal obtained the use of known concentrations of human AChR mRNA in independent experiments. Numbers and rs on the axes mark the position of each cDNA. (A) High-sensitivity fluorescein scan after hybridization fluorescein-labeled cDNA derived from wild-type plants. (B) Same array as in (A) but scanned at rate sensitivity. (C and D) A single array was probed with a 1:1 mixture of fluorescein-labeled cDNA wild-type plants and lissamine-labeled cDNA from HAT4-transgenic plants. The single array was scanned successively to detect the fluorescein fluorescence corresponding to mRNA from wild-type :s (C) and the lissamine fluorescence corresponding to mRNA from HAT4-transgenic plants (D). (E F) A single array was probed with a 1:1 mixture of fluorescein-labeled cDNA from root tissue and mine-labeled cDNA from leaf tissue. The single array was then scanned successively to detect the escein fluorescence corresponding to mRNAs expressed in roots (E) and the lissamine fluorescence sponding to mRNAs expressed in leaves (F).

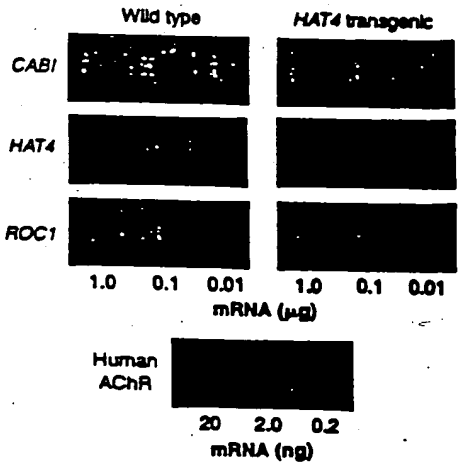


Fig. 2. Gene expression monitored with RNA (Northern) blot analysis. Designated amounts of mRNA from wild-type and *HAT4*-transgenic plants were spotted onto nylon membranes and probed with the cDNAs indicated. Purified human AChR mRNA was used for calibration.

and D, and Table 2). Hybridization of fluorescein-labeled glucocorticoid receptor cDNA (Fig. 1C) and lissamine-labeled *TRP4* cDNA (Fig. 1D) verified the presence of the negative control targets and the lack of optical cross talk between the two fluorophores.

To explore a more complex alteration in expression patterns, we performed a second two-color hybridization experiment with fluorescein- and lissamine-labeled probes prepared from root and leaf mRNA, respectively. The scanning sensitivities for the two fluorophores were normalized by matching the signals resulting from AChR mRNA, which was added to both cDNA synthesis reactions at a dilution of 1:1000 (Fig. 1, E and F). A comparison of the scans revealed widespread differences in gene expression between root and leaf tissue (Fig. 1, E and F). The mRNA from the light-regulated *CAB1* gene was ~500-fold more abundant in leaf (Fig. 1F) than in root tissue (Fig. 1E). The expression of 26 other genes differed between root and leaf tissue by more than a factor of 5 (Fig. 1, E and F).

The *HAT4*-transgenic line we examined has elongated hypocotyls, early flowering, poor germination, and altered pigmentation (*8*). Although changes in expression were observed for *HAT4*, large changes in expression were not observed for any of the other 44 genes we examined. This was somewhat surprising, particularly because comparative analysis of leaf and root tissue identified 27 differentially expressed genes. Analysis of an expanded set of genes may be required to identify genes whose expression changes upon *HAT4* overexpression; alternatively, a comparison of mRNA populations from specific tissues of wild-type and *HAT4*-transgenic plants may allow identification of downstream genes.

At the current density of robotic printing, it is feasible to scale up the fabrication process to produce arrays containing 20,000 cDNA targets. At this density, a single array would be sufficient to provide gene-specific targets encompassing nearly the entire repertoire of expressed genes in the *Arabidopsis* genome (*2*). The availability of 20,274 ESTs from *Arabidopsis* (*1*, *9*) would provide a rich source of templates for such studies.

The estimated 100,000 genes in the human genome (*10*) exceeds the number of *Arabidopsis* genes by a factor of 5 (*2*). This modest increase in complexity suggests that similar cDNA microarrays, prepared from the rapidly growing repertoire of human ESTs (*1*), could be used to determine the expression patterns of tens of thousands of human genes in diverse cell types. Coupling an amplification strategy to the reverse transcription reaction (*11*) could make it feasible to monitor expression even in minute tissue samples. A wide variety of acute and chronic physiological and pathological conditions might lead to characteristic changes in the patterns of gene expression in peripheral blood cells or other easily sampled tissues. In concert with cDNA microarrays for monitoring complex expression patterns, these tissues might therefore serve as sensitive in vivo sensors for clinical diagnosis. Microarrays of cDNAs could thus provide a useful link between human gene sequences and clinical medicine.

Table 1. Sequences contained on the cDNA microarray. Shown is the position, the known or putative function, and the accession number of each cDNA in the microarray (Fig. 1). All but three of the ESTs used in this study matched a sequence in the database. NADH, reduced form of nicotinamide adenine dinucleotide; ATPase, adenosine triphosphatase; GTP, guanosine triphosphate.

Position	cDNA	Function	Accession number
a1, 2	AChR	Human AChR	*
a3, 4	EST3	Actin	H36236
a5, 6	EST6	NADH dehydrogenase	Z27010
a7, 8	*AAC1*	Actin 1	M20016
a9, 10	EST12	Unknown	U36594†
a11, 12	EST13	Actin	T45783
b1, 2	*CAB1*	Chlorophyll a/b binding	M85150
b3, 4	EST17	Phosphoglycerate kinase	T44490
b5, 6	*GA4*	Gibberellic acid biosynthesis	L37126
b7, 8	EST19	Unknown	U36595†
b9, 10	*GBF-1*	G-box binding factor 1	X63894
b11, 12	EST23	Elongation factor	X52256
c1, 2	EST29	Aldolase	T04477
c3, 4	*GBF-2*	G-box binding factor 2	X63895
c5, 6	EST34	Chloroplast protease	R87034
c7, 8	EST35	Unknown	T14152
c9, 10	EST41	Catalase	T22720
c11, 12	rGR	Rat glucocorticoid receptor	M14053
d1, 2	EST42	Unknown	U36596†
d3, 4	EST45	ATPase	J04185
d5, 6	*HAT1*	Homeobox-leucine zipper 1	U09332
d7, 8	EST46	Light harvesting complex	T04063
d9, 10	EST49	Unknown	T76267
d11, 12	*HAT2*	Homeobox-leucine zipper 2	U09335
e1, 2	*HAT4*	Homeobox-leucine zipper 4	M90394
e3, 4	EST50	Phosphoribulokinase	T04344
e5, 6	*HAT5*	Homeobox-leucine zipper 5	M90416
7, 8	EST51	Unknown	Z33675
e9, 10	*HAT22*	Homeobox-leucine zipper 22	U09336
e11, 12	EST52	Oxygen evolving	T21749
f1, 2	EST59	Unknown	Z34607
f3, 4	*KNAT1*	*Knotted*-like homeobox 1	U14174
f5, 6	EST60	RuBisCO small subunit	X14564
f7, 8	EST69	Translation elongation factor	T42799
f9, 10	*PPH1*	Protein phosphatase 1	U34803
f11, 12	EST70	Unknown	T44621
g1, 2	EST75	Chloroplast protease	T43698
g3, 4	EST78	Unknown	R65481
g5, 6	*ROC1*	Cyclophilin	L14844
g7, 8	EST82	GTP binding	X59152
g9, 10	EST83	Unknown	Z33795
g11, 12	EST84	Unknown	T45278
h1, 2	EST91	Unknown	T13832
h3, 4	EST96	Unknown	R64816
h5, 6	*SAR1*	Synaptobrevin	M90418
h7, 8	EST100	Light harvesting complex	Z18205
h9, 10	EST103	Light harvesting complex	X03909
h11, 12	*TRP4*	Yeast tryptophan biosynthesis	X04273

*Proprietary sequence of Stratagene (La Jolla, California). †No match in the database; novel EST.

Table 2. Gene expression monitoring by microarray and RNA blot analyses; tg, *HAT4*-transgenic. See Table 1 for additional gene information. Expression levels (w/w) were calibrated with the use of known amounts of human AChR mRNA. Values for the microarray were determined from microarray scans (Fig. 1); values for the RNA blot were determined from RNA blots (Fig. 2).

Gene	Expression level (w/w)	
	Microarray	RNA blot
CAB1	1:48	1:83
CAB1 (tg)	1:120	1:150
HAT4	1:8300	1:8300
HAT4 (tg)	1:150	1:210
ROC1	1:1200	1:1800
ROC1 (tg)	1:260	1:1300

EFERENCES AND NOTES

rrent EST database (dbEST release 091495) e National Center for Biotechnology Informa- thesda, MD) contains a total of 322,225 en- ıcluding 255,645 from the human genome .044 from *Arabidopsis*. Access is available via rld Wide Web (http://www.ncbi.nlm.nih.gov). leyerowitz and R. E. Pruitt, *Science* 229, 1214 R. E. Pruitt and E. M. Meyerowitz, *J. Mol. Biol.* 9 (1986); I. Hwang *et al.*, *Plant J.* 1, 367 (1991); s *et al.*, *Plant Mol. Biol.* 24, 685 (1994); L. Le *t al.*, *Mol. Gen. Genet.* 245, 390 (1994).

on, thesis, Stanford University (1995); ——— O. Brown, in preparation. Microarrays were ed on poly-L-lysine–coated microscope Sigma) with a custom-built arraying machine ith one printing tip. The tip loaded 1 μl of PCR t (0.5 mg/ml) from 96-well microtiter plates posited ~0.005 μl per slide on 40 slides at a) of 500 μm. The printed slides were rehydrat- 2 hours in a humid chamber, snap-dried at for 1 min, rinsed in 0.1% SDS, and treated 05% succinic anhydride prepared in buffer ing of 50% 1-methyl-2-pyrrolidinone and onic acid. The cDNA on the slides was dena- distilled water for 2 min at 90°C immediately use. Microarrays were scanned with a laser ent scanner that contained a computer-con- (Y stage and a microscope objective. A mixed ultiline laser allowed sequential excitation of fluorophores. Emitted light was split accord- avelength and detected with two photomul- ibes. Signals were read into a PC with the use -bit analog-to-digital board. Additional details array fabrication and use may be obtained by of e-mail (pbrown@cmgm.stanford.edu).

usubel *et al.*, Eds., *Current Protocols in Mo- Biology* (Greene & Wiley Interscience, New 994), pp. 4.3.1–4.3.4.

enylated [poly(A)$^+$] mRNA was prepared from tA with the use of Oligotex-dT resin (Qiagen). e transcription (RT) reactions were carried out trataScript RT-PCR kit (Stratagene) modified ws: 50-μl reactions contained 0.1 μg/μl of ipsis mRNA, 0.1 ng/μl of human AChR 0.05 μg/μl of oligo(dT) (21-mer), 1× first buffer, 0.03 U/μl of ribonuclease block, 500 oxyadenosine triphosphate (dATP), 500 μM uanosine triphosphate, 500 μM dTTP, 40 oxycytosine triphosphate (dCTP), 40 μM flu- n-12-dCTP (or tissamine-5-dCTP), and 0.03 StrataScript reverse transcriptase. Reactions cubated for 60 min at 37°C, precipitated with , and resuspended in 10 μl of TE (10 mM tris- 1 mM EDTA, pH 8.0). Samples were then for 3 min at 94°C and chilled on ice. The RNA graded by adding 0.25 μl of 10 N NaOH 1 by a 10-min incubation at 37°C. The sam- re neutralized by addition of 2.5 μl of 1 M)H 8.0) and 0.25 μl of 10 N HCl and precip- ith ethanol. Pellets were washed with 70% dried to completion in a speedvac, resus- in 10 μl of H_2O, and reduced to 3.0 μl in a ac. Fluorescent nucleotide analogs were ob- om New England Nuclear (DuPont).

ation reactions contained 1.0 μl of fluorescent nthesis product (5) and 1.0 μl of hybridization 10× saline sodium citrate (SSC) and 0.2% he 2.0-μl probe mixtures were aliquoted onto oarray surface and covered with cover slips round). Arrays were transferred to a hybrid- chamber (3) and incubated for 18 hours at rrays were washed for 5 min at room temper- 5°C) in low-stringency wash buffer (1× SSC % SDS), then for 10 min at room temperature tringency wash buffer (0.1× SSC and 0.1% rrays were scanned in 0.1× SSC with the use escence laser-scanning device (3).

s of poly(A)$^+$ mRNA (4, 5) were spotted onto embranes (Nytran) and crosslinked with ul- light with the use of a Stratalinker 1800 ene). Probes were prepared by random with the use of a Prime-It II kit (Stratagene) in ence of [^{32}P]dATP. Hybridizations were car- according to the instructions of the manu- facturer. Quantitation was performed on a Phosphorimager (Molecular Dynamics).

8. M. Schena and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* 89, 3894 (1992); M. Schena, A. M. Lloyd, R. W. Davis, *Genes Dev.* 7, 367 (1993); M. Schena and R. W. Davis, *Proc. Natl. Acad. Sci. U.S.A.* 91, 8393 (1994).
9. H. Hofte *et al.*, *Plant J.* 4, 1051 (1993); T. Newman *et al.*, *Plant Physiol.* 106, 1241 (1994).
10. N. E. Morton, *Proc. Natl. Acad. Sci. U.S.A.* 88, 7474 (1991); E. D. Green and R. H. Waterston, *J. Am. Med. Assoc.* 266, 1966 (1991); C. Betsame-Chantelot, *Cell* 70, 1059 (1992); D. R. Cox *et al.*, *Science* 265, 2031 (1994).
11. E. S. Kawasaki *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 85, 5698 (1988).
12. The laser fluorescent scanner was designed and fabricated in collaboration with S. Smith of Stanford University. Scanner and analysis software was developed by R. X. Xia. The succinic anhydride reaction was suggested by J. Mulligan and J. Van Ness of Darwin Molecular Corporation. Thanks to S. Theologis, C. Somerville, K. Yamamoto, and members of the laboratories of R.W.D. and P.O.B. for critical comments. Supported by the Howard Hughes Medical Institute and by grants from NIH (R21HG00450) (P.O.B.) and F37AG00195 (R.W.D.)] and from NSF (MCB9106011) (R.W.D.) and by an NSF graduate fellowship (D.S.). P.O.B. is an assistant investigator of the Howard Hughes Medical Institute.

11 August 1995; accepted 22 September 1995

Gene Therapy in Peripheral Blood Lymphocytes and Bone Marrow for ADA$^-$ Immunodeficient Patients

Claudio Bordignon,* Luigi D. Notarangelo, Nadia Nobili, Giuliana Ferrari, Giulia Casorati, Paola Panina, Evelina Mazzolari, Daniela Maggioni, Claudia Rossi, Paolo Servida, Alberto G. Ugazio, Fulvio Mavilio

Adenosine deaminase (ADA) deficiency results in severe combined immunodeficiency, the first genetic disorder treated by gene therapy. Two different retroviral vectors were used to transfer ex vivo the human ADA minigene into bone marrow cells and peripheral blood lymphocytes from two patients undergoing exogenous enzyme replacement therapy. After 2 years of treatment, long-term survival of T and B lymphocytes, marrow cells, and granulocytes expressing the transferred ADA gene was demonstrated and resulted in normalization of the immune repertoire and restoration of cellular and humoral immunity. After discontinuation of treatment, T lymphocytes, derived from transduced peripheral blood lymphocytes, were progressively replaced by marrow-derived T cells in both patients. These results indicate successful gene transfer into long-lasting progenitor cells, producing a functional multilineage progeny.

Severe combined immunodeficiency associated with inherited deficiency of ADA (*1*) is usually fatal unless affected children are kept in protective isolation or the immune system is reconstituted by bone marrow transplantation from a human leukocyte antigen (HLA)–identical sibling donor (*2*). This is the therapy of choice, although it is available only for a minority of patients. In recent years, other forms of therapy have been developed, including transplants from haploidentical donors (*3*, *4*), exogenous enzyme replacement (*5*), and somatic-cell gene therapy (*6*–*9*).

We previously reported a preclinical model in which ADA gene transfer and expression successfully restored immune functions in human ADA-deficient (ADA$^-$) peripheral blood lymphocytes (PBLs) in immunodeficient mice in vivo (*10*, *11*). On the basis of these preclinical results, the clinical application of gene therapy for the treatment of ADA$^-$ SCID (severe combined immunodeficiency disease) patients who previously failed exogenous enzyme replacement therapy was approved by our Institutional Ethical Committees and by the Italian National Committee for Bioethics (*12*). In addition to evaluating the safety and efficacy of the gene therapy procedure, the aim of the study was to define the relative role of PBLs and hematopoietic stem cells in the long-term reconstitution of immune functions after retroviral vector–mediated ADA gene transfer. For this purpose, two structurally identical vectors expressing the human ADA complementary DNA (cDNA), distinguishable by the presence of alternative restriction sites in a nonfunctional region of the viral long-terminal repeat (LTR), were used to transduce PBLs and bone marrow (BM) cells independently. This procedure allowed identification of the origin of

C. Bordignon, N. Nobili, G. Ferrari, D. Maggioni, C. Rossi, P. Servida, F. Mavilio, Telethon Gene Therapy Program for Genetic Diseases, DIBIT, Istituto Scientifico H. S. Raffaele, Milan, Italy.
L. D. Notarangelo, E. Mazzolari, A. G. Ugazio, Department of Pediatrics, University of Brescia Medical School, Brescia, Italy.
G. Casorati, Unità di Immunochimica, DIBIT, Istituto Scientifico H. S. Raffaele, Milan, Italy.
P. Panina, Roche Milano Ricerche, Milan, Italy.
*To whom correspondence should be addressed.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : G01N 33/543, 33/68

A1

(11) International Publication Number: WO 95/35505

(43) International Publication Date: 28 December 1995 (28.12.95)

(21) International Application Number: PCT/US95/07659

(22) International Filing Date: 16 June 1995 (16.06.95)

(30) Priority Data:

08/261,388	17 June 1994 (17.06.94)	US
08/477,809	7 June 1995 (07.06.95)	US

(71) Applicant: THE BOARD OF TRUSTEES OF THE LELAND STANFORD JUNIOR UNIVERSITY [US/US]; Stanford, CA 94305 (US).

(72) Inventors: SHALON, Tidhar, Dari; 364 Fletcher Drive, Atherton, CA 94027 (US). BROWN, Patrick, O.; 76 Peter Coutts Circle, Stanford, CA 94305 (US).

(74) Agent: DEHLINGER, Peter, J.; Dehlinger & Associates, P.O. Box 60850, Palo Alto, CA 94306-1546 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: METHOD AND APPARATUS FOR FABRICATING MICROARRAYS OF BIOLOGICAL SAMPLES

(57) Abstract

A method and apparatus for forming microarrays of biological samples on a support are disclosed. The method involves dispensing a known volume of a reagent at each of a selected array position, by tapping a capillary dispenser on the support under conditions effective to draw a defined volume of liquid onto the support. The apparatus is designed to produce a microarray of such regions in an automated fashion.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**GB**	United Kingdom	**MR**	Mauritania
AU	Australia	**GE**	Georgia	**MW**	Malawi
BB	Barbados	**GN**	Guinea	**NE**	Niger
BE	Belgium	**GR**	Greece	**NL**	Netherlands
BF	Burkina Faso	**HU**	Hungary	**NO**	Norway
BG	Bulgaria	**IE**	Ireland	**NZ**	New Zealand
BJ	Benin	**IT**	Italy	**PL**	Poland
BR	Brazil	**JP**	Japan	**PT**	Portugal
BY	Belarus	**KE**	Kenya	**RO**	Romania
CA	Canada	**KG**	Kyrgystan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SI**	Slovenia
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SK**	Slovakia
CM	Cameroon	**LI**	Liechtenstein	**SN**	Senegal
CN	China	**LK**	Sri Lanka	**TD**	Chad
CS	Czechoslovakia	**LU**	Luxembourg	**TG**	Togo
CZ	Czech Republic	**LV**	Latvia	**TJ**	Tajikistan
DE	Germany	**MC**	Monaco	**TT**	Trinidad and Tobago
DK	Denmark	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**US**	United States of America
FI	Finland	**ML**	Mali	**UZ**	Uzbekistan
FR	France	**MN**	Mongolia	**VN**	Viet Nam
GA	Gabon				

METHOD AND APPARATUS FOR FABRICATING MICROARRAYS OF BIOLOGICAL SAMPLES

Field of the Invention

This invention relates to a method and apparatus for fabricating microarrays of biological samples for large scale screening assays, such as arrays of DNA samples to be used in DNA hybridization assays for genetic research and diagnostic applications.

References

Abouzied, *et al.*, *Journal of AOAC International* 77(2):495-500 (1994).

Bohlander, *et al.*, *Genomics* 13:1322-1324 (1992).

Drmanac, *et al.*, *Science* 260:1649-1652 (1993).

Fodor, *et al.*, *Science* 251:767-773 (1991).

Khrapko, *et al.*, *DNA Sequence* 1:375-388 (1991).

Kuriyama, *et al.*, AN ISFET BIOSENSOR, APPLIED BIOSENSORS (Donald Wise, Ed.), Butterworths, pp. 93-114 (1989).

Lehrach, *et al.*, HYBRIDIZATION FINGERPRINTING IN GENOME MAPPING AND SEQUENCING, GENOME ANALYSIS, VOL 1 (Davies and Tilgham, Eds.), Cold Spring Harbor Press, pp. 39-81 (1990).

Maniatis, *et al.*, MOLECULAR CLONING, A LABORATORY MANUAL, Cold Spring Harbor Press (1989).

Nelson, *et al.*, *Nature Genetics* 4:11-18 (1993).

Pirrung, *et al.*, U.S. Patent No. 5,143,854 (1992).

Riles, *et al.*, *Genetics* 134:81-150 (1993).

Schena, M. *et al.*, *Proc. Nat. Acad. Sci. USA* 89:3894-3898 (1992).

Southern, *et al.*, *Genomics* 13:1008-1017 (1992).

Background of the Invention

A variety of methods are currently available for making arrays of biological macromolecules, such as arrays of nucleic acid molecules or proteins. One method for making ordered arrays of DNA on a porous membrane is a "dot blot" approach. In this method, a vacuum manifold transfers a plurality, *e.g.*, 96, aqueous samples of DNA from 3 millimeter diameter wells to a porous membrane. A common variant of this procedure is a "slot-blot" method in which the wells have highly-elongated oval shapes.

The DNA is immobilized on the porous membrane by baking the membrane or exposing it to UV radiation. This is a manual procedure practical for making one array at a time and usually limited to 96 samples per array. "Dot-blot" procedures are therefore inadequate for applications in which many thousand samples must be determined.

A more efficient technique employed for making ordered arrays of genomic fragments uses an array of pins dipped into the wells, *e.g.*, the 96 wells of a microtitre plate, for transferring an array of samples to a substrate, such as a porous membrane. One array includes pins that are designed to spot a membrane in a staggered fashion, for creating an array of 9216 spots in a 22 × 22 cm area (Lehrach, *et al.*, 1990). A limitation with this approach is that the volume of DNA spotted in each pixel of each array is highly variable.

In addition, the number of arrays that can be made with each dipping is usually quite small.

An alternate method of creating ordered arrays of nucleic acid sequences is described by Pirrung, *et al.* (1992), and also by Fodor, *et al.* (1991). The method involves synthesizing different nucleic acid sequences at different discrete regions of a support. This method employs elaborate synthetic schemes, and is generally limited to relatively short nucleic acid sample, *e.g.*, less than 20 bases. A related method has been described by Southern, *et al.* (1992).

Khrapko, *et al.* (1991) describes a method of making an oligonucleotide matrix by spotting DNA onto a thin layer of polyacrylamide. The spotting is done manually with a micropipette.

None of the methods or devices described in the prior art are designed for mass fabrication of microarrays characterized by (i) a large number of micro-sized assay regions separated by a distance of 50-200 microns or less, and (ii) a well-defined amount, typically in the picomole range, of analyte associated with each region of the array.

Furthermore, current technology is directed at performing such assays one at a time to a single array of DNA molecules. For example, the most common method for performing DNA hybridizations to arrays spotted onto porous membrane involves sealing the membrane in a plastic bag (Maniatas, *et al.*, 1989) or a rotating glass cylinder (Robbins Scientific) with the labeled hybridization probe inside the sealed chamber. For arrays made on non-porous surfaces, such as a microscope slide, each array is incubated with the labeled hybridization probe sealed under a coverslip. These techniques require a separate sealed chamber for

each array which makes the screening and handling of many such arrays inconvenient and time intensive.

Abouzied, *et al.* (1994) describes a method of printing horizontal lines of antibodies on a nitrocellulose membrane and separating regions of the membrane with vertical stripes of a hydrophobic material. Each vertical stripe is then reacted with a different antigen and the reaction between the immobilized antibody and an antigen is detected using a standard ELISA colorimetric technique. Abouzied's technique makes it possible to screen many one-dimensional arrays simultaneously on a single sheet of nitrocellulose. Abouzied makes the nitrocellulose somewhat hydrophobic using a line drawn with PAP Pen (Research Products International). However Abouzied does not describe a technology that is capable of completely sealing the pores of the nitrocellulose. The pores of the nitrocellulose are still physically open and so the assay reagents can leak through the hydrophobic barrier during extended high temperature incubations or in the presence of detergents which makes the Abouzied technique unacceptable for DNA hybridization assays.

Porous membranes with printed patterns of hydrophilic/hydrophobic regions exist for applications such as ordered arrays of bacteria colonies. QA Life Sciences (San Diego CA) makes such a membrane with a grid pattern printed on it. However, this membrane has the same disadvantage as the Abouzied technique since reagents can still flow between the gridded arrays making them unusable for separate DNA hybridization assays.

Pall Corporation make a 96-well plate with a porous filter heat sealed to the bottom of the plate. These plates are capable of containing different

reagents in each well without cross-contamination. However, each well is intended to hold only one target element whereas the invention described here makes a microarray of many biomolecules in each subdivided region of the solid support. Furthermore, the 96 well plates are at least 1 cm thick and prevent the use of the device for many colorimetric, fluorescent and radioactive detection formats which require that the membrane lie flat against the detection surface. The invention described here requires no further processing after the assay step since the barriers elements are shallow and do not interfere with the detection step thereby greatly increasing convenience.

Hyseq Corporation has described a method of making an "array of arrays" on a non-porous solid support for use with their sequencing by hybridization technique. The method described by Hyseq involves modifying the chemistry of the solid support material to form a hydrophobic grid pattern where each subdivided region contains a microarray of biomolecules. Hyseq's flat hydrophobic pattern does not make use of physical blocking as an additional means of preventing cross contamination.

Summary of the Invention

The invention includes, in one aspect, a method of forming a microarray of analyte-assay regions on a solid support, where each region in the array has a known amount of a selected, analyte-specific reagent. The method involves first loading a solution of a selected analyte-specific reagent in a reagent-dispensing device having an elongate capillary channel (i) formed by spaced-apart, coextensive elongate members, (ii) adapted to hold a quantity of the reagent solution and (iii) having a tip region at which aqueous

solution in the channel forms a meniscus. The channel is preferably formed by a pair of spaced-apart tapered elements.

The tip of the dispensing device is tapped against a solid support at a defined position on the support surface with an impulse effective to break the meniscus in the capillary channel deposit a selected volume of solution on the surface, preferably a selected volume in the range 0.01 to 100 nl. The two steps are repeated until the desired array is formed.

The method may be practiced in forming a plurality of such arrays, where the solution-depositing step is are applied to a selected position on each of a plurality of solid supports at each repeat cycle.

The dispensing device may be loaded with a new solution, by the steps of (i) dipping the capillary channel of the device in a wash solution, (ii) removing wash solution drawn into the capillary channel, and (iii) dipping the capillary channel into the new reagent solution.

Also included in the invention is an automated apparatus for forming a microarray of analyte-assay regions on a plurality of solid supports, where each region in the array has a known amount of a selected, analyte-specific reagent. The apparatus has a holder for holding, at known positions, a plurality of planar supports, and a reagent dispensing device of the type described above.

The apparatus further includes positioning structure for positioning the dispensing device at a selected array position with respect to a support in said holder, and dispensing structure for moving the dispensing device into tapping engagement against a support with a selected impulse effective to deposit a

selected volume on the support, .g., a selected volume in the volume range 0.01 to 100 nl.

The positioning and dispensing structures are controlled by a control unit in the apparatus. The unit operates to (i) place the dispensing device at a loading station, (ii) move the capillary channel in the device into a selected reagent at the loading station, to load the dispensing device with the reagent, and (iii) dispense the reagent at a defined array position on each of the supports on said holder. The unit may further operate, at the end of a dispensing cycle, to wash the dispensing device by (i) placing the dispensing device at a washing station, (ii) moving the capillary channel in the device into a wash fluid, to load the dispensing device with the fluid, and (iii) remove the wash fluid prior to loading the dispensing device with a fresh selected reagent.

The dispensing device in the apparatus may be one of a plurality of such devices which are carried on the arm for dispensing different analyte assay reagents at selected spaced array positions.

In another aspect, the invention includes a substrate with a surface having a microarray of at least 10^3 distinct polynucleotide or polypeptide biopolymers in a surface area of less than about 1 cm^2. Each distinct biopolymer (i) is disposed at a separate, defined position in said array, (ii) has a length of at least 50 subunits, and (iii) is present in a defined amount between about 0.1 femtomoles and 100 nanomoles.

In one embodiment, the surface is glass slide surface coated with a polycationic polymer, such as polylysine, and the biopolymers are polynucleotides. In another embodiment, the substrate has a water-impermeable backing, a water-permeable film formed on

the backing, and a grid formed on the film. The grid is composed of intersecting water-impervious grid elements extending from said backing to positions raised above the surface of said film, and partitions the film into a plurality of water-impervious cells. A biopolymer array is formed within each well.

More generally, there is provided a substrate for use in detecting binding of labeled polynucleotides to one or more of a plurality different-sequence, immobilized polynucleotides. The substrate includes, in one aspect, a glass support, a coating of a polycationic polymer, such as polylysine, on said surface of the support, and an array of distinct polynucleotides electrostatically bound non-covalently to said coating, where each distinct biopolymer is disposed at a separate, defined position in a surface array of polynucleotides.

In another aspect, the substrate includes a water-impermeable backing, a water-permeable film formed on the backing, and a grid formed on the film, where the grid is composed of intersecting water-impervious grid elements extending from the backing to positions raised above the surface of the film, forming a plurality of cells. A biopolymer array is formed within each cell.

Also forming part of the invention is a method of detecting differential expression of each of a plurality of genes in a first cell type, with respect to expression of the same genes in a second cell type. In practicing the method, there is first produced fluorescent-labeled cDNA's from mRNA's isolated from the two cells types, where the cDNA'S from the first and second cells are labeled with first and second different fluorescent reporters.

A mixture of the labeled cDNA's from the two cell types is added to an array of polynucleotides

representing a plurality of known genes derived from the two cell types, under conditions that result in hybridization of the cDNA's to complementary-sequence polynucleotides in the array. The array is then examined by fluorescence under fluorescence excitation conditions in which (i) polynucleotides in the array that are hybridized predominantly to cDNA's derived from one of the first and second cell types give a distinct first or second fluorescence emission color, respectively, and (ii) polynucleotides in the array that are hybridized to substantially equal numbers of cDNA's derived from the first and second cell types give a distinct combined fluorescence emission color, respectively. The relative expression of known genes in the two cell types can then be determined by the observed fluorescence emission color of each spot.

These and other objects and features of the invention will become more fully apparent when the following detailed description of the invention is read in conjunction with the accompanying figures.

Brief Description of the Drawings

Fig. 1 is a side view of a reagent-dispensing device having a open-capillary dispensing head constructed for use in one embodiment of the invention;

Figs. 2A-2C illustrate steps in the delivery of a fixed-volume bead on a hydrophobic surface employing the dispensing head from Fig. 1, in accordance with one embodiment of the method of the invention;

Fig. 3 shows a portion of a two-dimensional array of analyte-assay regions constructed according to the method of the invention;

Fig. 4 is a planar view showing components of an automated apparatus for forming arrays in accordance with the inventi n.

Fig. 5 shows a fluorescent image of an actual 20 × 20 array of 400 fluorescently-labeled DNA samples immobilized on a poly-l-lysine coated slide, where the total area covered by the 400 element array is 16 square millimeters;

Fig. 6 is a fluorescent image of a 1.8 cm × 1.8 cm microarray containing lambda clones with yeast inserts, the fluorescent signal arising from the hybridization to the array with approximately half the yeast genome labeled with a green fluorophore and the other half with a red fluorophore;

Fig. 7 shows the translation of the hybridization image of Fig. 6 into a karyotype of the yeast genome, where the elements of Fig.-6 microarray contain yeast DNA sequences that have been previously physically mapped in the yeast genome;

Fig. 8 show a fluorescent image of a 0.5 cm × 0.5 cm microarray of 24 cDNA clones, where the microarray was hybridized simultaneously with total cDNA from wild type *Arabidopsis* plant labeled with a green fluorophore and total cDNA from a transgenic *Arabidopsis* plant labeled with a red fluorophore, and the arrow points to the cDNA clone representing the gene introduced into the transgenic *Arabidopsis* plant;

Fig. 9 shows a plan view of substrate having an array of cells formed by barrier elements in the form of a grid;

Fig. 10 shows an enlarged plan view of one of the cells in the substrate in Fig. 9, showing an array of polynucleotide regions in the cell;

Fig. 11 is an enlarged sectional view of the substrate in Fig. 9, taken along a section line in that figure; and

Fig. 12 is a scann d image of a 3 cm × 3 cm nitrocellulose solid support containing four identical

arrays of M13 clones in each of four quadrants, where each quadrant was hybridized simultaneously to a different oligonucleotide using an open face hybridization method.

Detailed Description of the Invention

I. Definitions

Unless indicated otherwise, the terms defined below have the following meanings:

"Ligand" refers to one member of a ligand/anti-ligand binding pair. The ligand may be, for example, one of the nucleic acid strands in a complementary, hybridized nucleic acid duplex binding pair; an effector molecule in an effector/receptor binding pair; or an antigen in an antigen/antibody or antigen/antibody fragment binding pair.

"Antiligand" refers to the opposite member of a ligand/anti-ligand binding pair. The antiligand may be the other of the nucleic acid strands in a complementary, hybridized nucleic acid duplex binding pair; the receptor molecule in an effector/receptor binding pair; or an antibody or antibody fragment molecule in antigen/antibody or antigen/antibody fragment binding pair, respectively.

"Analyte" or "analyte molecule" refers to a molecule, typically a macromolecule, such as a polynucleotide or polypeptide, whose presence, amount, and/or identity are to be determined. The analyte is one member of a ligand/anti-ligand pair.

"Analyte-specific assay reagent" refers to a molecule effective to bind specifically to an analyte molecule. The reagent is the opposite member of a ligand/anti-ligand binding pair.

An "array f regions on a solid support" is a linear or two-dimensional array of preferably discrete

regions, each having a finite area, formed on the surface of a solid support.

A "microarray" is an array of regions having a density of discrete regions of at least about 100/cm^2, and preferably at least about 1000/cm^2. The regions in a microarray have typical dimensions, e.g., diameters, in the range of between about 10-250 μm, and are separated from other regions in the array by about the same distance.

A support surface is "hydrophobic" if a aqueous-medium droplet applied to the surface does not spread out substantially beyond the area size of the applied droplet. That is, the surface acts to prevent spreading of the droplet applied to the surface by hydrophobic interaction with the droplet.

A "meniscus" means a concave or convex surface that forms on the bottom of a liquid in a channel as a result of the surface tension of the liquid.

"Distinct biopolymers", as applied to the biopolymers forming a microarray, means an array member which is distinct from other array members on the basis of a different biopolymer sequence, and/or different concentrations of the same or distinct biopolymers, and/or different mixtures of distinct or different-concentration biopolymers. Thus an array of "distinct polynucleotides" means an array containing, as its members, (i) distinct polynucleotides, which may have a defined amount in each member, (ii) different, graded concentrations of given-sequence polynucleotides, and/or (iii) different-composition mixtures of two or more distinct polynucleotides.

"Cell type" means a cell from a given source, e.g., a tissue, or organ, or a cell in a given state of

differentiation, or a cell associated with a given pathology or genetic makeup.

II. Method of Microarray Formation

This section describes a method of forming a microarray of analyte-assay regions on a solid support or substrate, where each region in the array has a known amount of a selected, analyte-specific reagent.

Fig. 1 illustrates, in a partially schematic view, a reagent-dispensing device 10 useful in practicing the method. The device generally includes a reagent dispenser 12 having an elongate open capillary channel 14 adapted to hold a quantity of the reagent solution, such as indicated at 16, as will be described below. The capillary channel is formed by a pair of spaced-apart, coextensive, elongate members 12a, 12b which are tapered toward one another and converge at a tip or tip region 18 at the lower end of the channel. More generally, the open channel is formed by at least two elongate, spaced-apart members adapted to hold a quantity of reagent solutions and having a tip region at which aqueous solution in the channel forms a meniscus, such as the concave meniscus illustrated at 20 in Fig. 2A. The advantages of the open channel construction of the dispenser are discussed below.

With continued reference to Fig. 1, the dispenser device also includes structure for moving the dispenser rapidly toward and away from a support surface, for effecting deposition of a known amount of solution in the dispenser on a support, as will be described below with reference to Figs. 2A-2C. In the embodiment shown, this structure includes a solenoid 22 which is activatable to draw a solenoid piston 24 rapidly downwardly, then release the piston, *e.g.*, under spring bias, to a normal, raised position, as shown. The

dispenser is carried on the piston by a connecting member 26, as shown. The just-described moving structure is also referred to herein as dispensing means for moving the dispenser into engagement with a solid support, for dispensing a known volume of fluid on the support.

The dispensing device just described is carried on an arm 28 that may be moved either linearly or in an x-y plane to position the dispenser at a selected deposition position, as will be described.

Figs. 2A-2C illustrate the method of depositing a known amount of reagent solution in the just-described dispenser on the surface of a solid support, such as the support indicated at 30. The support is a polymer, glass, or other solid-material support having a surface indicated at 31.

In one general embodiment, the surface is a relatively hydrophilic, *i.e.*, wettable surface, such as a surface having native, bound or covalently attached charged groups. On such surface described below is a glass surface having an absorbed layer of a polycationic polymer, such as poly-l-lysine.

In another embodiment, the surface has or is formed to have a relatively hydrophobic character, *i.e.*, one that causes aqueous medium deposited on the surface to bead. A variety of known hydrophobic polymers, such as polystyrene, polypropylene, or polyethylene have desired hydrophobic properties, as do glass and a variety of lubricant or other hydrophobic films that may be applied to the support surface.

Initially, the dispenser is loaded with a selected analyte-specific reagent solution, such as by dipping the dispenser tip, after washing, into a solution of the reagent, and allowing filling by capillary flow into the dispenser channel. The dispenser is now moved

to a selected position with respect to a support surface, placing the dispenser tip directly above the support-surface position at which the reagent is to be deposited. This movement takes place with the dispenser tip in its raised position, as seen in Fig. 2A, where the tip is typically at least several 1-5 mm above the surface of the substrate.

With the dispenser so positioned, solenoid 22 is now activated to cause the dispenser tip to move rapidly toward and away from the substrate surface, making momentary contact with the surface, in effect, tapping the tip of the dispenser against the support surface. The tapping movement of the tip against the surface acts to break the liquid meniscus in the tip channel, bringing the liquid in the tip into contact with the support surface. This, in turn, produces a flowing of the liquid into the capillary space between the tip and the surface, acting to draw liquid out of the dispenser channel, as seen in Fig. 2B.

Fig. 2C shows flow of fluid from the tip onto the support surface, which in this case is a hydrophobic surface. The figure illustrates that liquid continues to flow from the dispenser onto the support surface until it forms a liquid bead 32. At a given bead size, *i.e.*, volume, the tendency of liquid to flow onto the surface will be balanced by the hydrophobic surface interaction of the bead with the support surface, which acts to limit the total bead area on the surface, and by the surface tension of the droplet, which tends toward a given bead curvature. At this point, a given bead volume will have formed, and continued contact of the dispenser tip with the bead, as the dispenser tip is being withdrawn, will have little or no effect on bead volume.

For liquid-dispensing on a more hydrophilic surface, the liquid will have less of a tendency to bead, and the dispensed volume will be more sensitive to the total dwell time of the dispenser tip in the immediate vicinity of the support surface, *e.g.*, the positions illustrated in Figs. 2B and 2C.

The desired deposition volume, *i.e.*, bead volume, formed by this method is preferably in the range 2 pl (picoliters) to 2 nl (nanoliters), although volumes as high as 100 nl or more may be dispensed. It will be appreciated that the selected dispensed volume will depend on (i) the "footprint" of the dispenser tip, *i.e.*, the size of the area spanned by the tip, (ii) the hydrophobicity of the support surface, and (iii) the time of contact with and rate of withdrawal of the tip from the support surface. In addition, bead size may be reduced by increasing the viscosity of the medium, effectively reducing the flow time of liquid from the dispenser onto the support surface. The drop size may be further constrained by depositing the drop in a hydrophilic region surrounded by a hydrophobic grid pattern on the support surface.

In a typical embodiment, the dispenser tip is tapped rapidly against the support surface, with a total residence time in contact with the support of less than about 1 msec, and a rate of upward travel from the surface of about 10 cm/sec.

Assuming that the bead that forms on contact with the surface is a hemispherical bead, with a diameter approximately equal to the width of the dispenser tip, as shown in Fig. 2C, the volume of the bead formed in relation to dispenser tip width (d) is given in Table 1 below. As seen, the volume of the bead ranges between 2 pl to 2 nl as the width size is increased from about 20 to 200 μm.

Table 1

d	Volume (nl)
20 μm	2×10^{-3}
50 μm	3.1×10^{-2}
100 μm	2.5×10^{-1}
200 μm	2

At a given tip size, bead volume can be reduced in a controlled fashion by increasing surface hydrophobicity, reducing time of contact of the tip with the surface, increasing rate of movement of the tip away from the surface, and/or increasing the viscosity of the medium. Once these parameters are fixed, a selected deposition volume in the desired pl to nl range can be achieved in a repeatable fashion.

After depositing a bead at one selected location on a support, the tip is typically moved to a corresponding position on a second support, a droplet is deposited at that position, and this process is repeated until a liquid droplet of the reagent has been deposited at a selected position on each of a plurality of supports.

The tip is then washed to remove the reagent liquid, filled with another reagent liquid and this reagent is now deposited at each another array position on each of the supports. In one embodiment, the tip is washed and refilled by the steps of (i) dipping the capillary channel of the device in a wash solution, (ii) removing wash solution drawn into the capillary channel, and (iii) dipping the capillary channel into the new reagent solution.

From the foreg ing, it will be appreciated that the tweezers-like, open-capillary dispenser tip

provides the advantages that (i) the open channel of the tip facilitates rapid, efficient washing and drying before reloading the tip with a new reagent, (ii) passive capillary action can load the sample directly from a standard microwell plate while retaining sufficient sample in the open capillary reservoir for the printing of numerous arrays, (iii) open capillaries are less prone to clogging than closed capillaries, and (iv) open capillaries do not require a perfectly faced bottom surface for fluid delivery.

A portion of a microarray 36 formed on the surface 38 of a solid support 40 in accordance with the method just described is shown in Fig. 3. The array is formed of a plurality of analyte-specific reagent regions, such as regions 42, where each region may include a different analyte-specific reagent. As indicated above, the diameter of each region is preferably between about 20-200 μm. The spacing between each region and its closest (non-diagonal) neighbor, measured from center-to-center (indicated at 44), is preferably in the range of about 20-400 μm. Thus, for example, an array having a center-to-center spacing of about 250 μm contains about 40 regions/cm or 1,600 regions/cm^2. After formation of the array, the support is treated to evaporate the liquid of the droplet forming each region, to leave a desired array of dried, relatively flat regions. This drying may be done by heating or under vacuum.

In some cases, it is desired to first rehydrate the droplets containing the analyte reagents to allow for more time for adsorption to the solid support. It is also possible to spot out the analyte reagents in a humid environment so that droplets do not dry until the arraying operation is complete.

III. Automated Apparatus for Forming Arrays

In another aspect, the invention includes an automated apparatus for forming an array of analyte-assay regions on a solid support, where each region in the array has a known amount of a selected, analyte-specific reagent.

The apparatus is shown in planar, and partially schematic view in Fig. 4. A dispenser device 72 in the apparatus has the basic construction described above with respect to Fig. 1, and includes a dispenser 74 having an open-capillary channel terminating at a tip, substantially as shown in Figs. 1 and 2A-2C.

The dispenser is mounted in the device for movement toward and away from a dispensing position at which the tip of the dispenser taps a support surface, to dispense a selected volume of reagent solution, as described above. This movement is effected by a solenoid 76 as described above. Solenoid 76 is under the control of a control unit 77 whose operation will be described below. The solenoid is also referred to herein as dispensing means for moving the device into tapping engagement with a support, when the device is positioned at a defined array position with respect to that support.

The dispenser device is carried on an arm 74 which is threadedly mounted on a worm screw 80 driven (rotated) in a desired direction by a stepper motor 82 also under the control of unit 77. At its left end in the figure screw 80 is carried in a sleeve 84 for rotation about the screw axis. At its other end, the screw is mounted to the drive shaft of the stepper motor, which in turn is carried on a sleeve 86. The dispenser device, worm screw, the two sleeves mounting the w rm screw, and the stepper motor used in moving the device in the "x" (horizontal) direction in the

figure form what is referred to here collectively as a displacement assembly 86.

The displacement assembly is constructed to produce precise, micro-range movement in the direction of the screw, *i.e.*, along an x axis in the figure. In one mode, the assembly functions to move the dispenser in x-axis increments having a selected distance in the range 5-25 μm. In another mode, the dispenser unit may be moved in precise x-axis increments of several microns or more, for positioning the dispenser at associated positions on adjacent supports, as will be described below.

The displacement assembly, in turn, is mounted for movement in the "y" (vertical) axis of the figure, for positioning the dispenser at a selected y axis position. The structure mounting the assembly includes a fixed rod 88 mounted rigidly between a pair of frame bars 90, 92, and a worm screw 94 mounted for rotation between a pair of frame bars 96, 98. The worm screw is driven (rotated) by a stepper motor 100 which operates under the control of unit 77. The motor is mounted on bar 96, as shown.

The structure just described, including worm screw 94 and motor 100, is constructed to produce precise, micro-range movement in the direction of the screw, *i.e.*, along an y axis in the figure. As above, the structure functions in one mode to move the dispenser in y-axis increments having a selected distance in the range 5-250 μm, and in a second mode, to move the dispenser in precise y-axis increments of several microns (μm) or more, for positioning the dispenser at associated positions on adjacent supports.

The displacement assembly and structure for moving this assembly in the y axis are referred to herein collectively as positioning means for positioning the

dispensing device at a selected array position with respect to a support.

A holder 102 in the apparatus functions to hold a plurality of supports, such as supports 104 on which the microarrays of regent regions are to be formed by the apparatus. The holder provides a number of recessed slots, such as slot 106, which receive the supports, and position them at precise selected positions with respect to the frame bars on which the dispenser moving means is mounted.

As noted above, the control unit in the device functions to actuate the two stepper motors and dispenser solenoid in a sequence designed for automated operation of the apparatus in forming a selected microarray of reagent regions on each of a plurality of supports.

The control unit is constructed, according to conventional microprocessor control principles, to provide appropriate signals to each of the solenoid and each of the stepper motors, in a given timed sequence and for appropriate signalling time. The construction of the unit, and the settings that are selected by the user to achieve a desired array pattern, will be understood from the following description of a typical apparatus operation.

Initially, one or more supports are placed in one or more slots in the holder. The dispenser is then moved to a position directly above a well (not shown) containing a solution of the first reagent to be dispensed on the support(s). The dispenser solenoid is actuated now to lower the dispenser tip into this well, causing the capillary channel in the dispenser to fill. Motors 82, 100 are now actuated to position the dispenser at a s lected array position at the first of the supports. Solenoid actuation of the dispenser is

then effective to dispens a selected-volume droplet of that reagent at this location. As noted above, this operation is effective to dispense a selected volume preferably between 2 pl and 2 nl of the reagent solution.

The dispenser is now moved to the corresponding position at an adjacent support and a similar volume of the solution is dispensed at this position. The process is repeated until the reagent has been dispensed at this preselected corresponding position on each of the supports.

Where it is desired to dispense a single reagent at more than two array positions on a support, the dispenser may be moved to different array positions at each support, before moving the dispenser to a new support, or solution can be dispensed at individual positions on each support, at one selected position, then the cycle repeated for each new array position.

To dispense the next reagent, the dispenser is positioned over a wash solution (not shown), and the dispenser tip is dipped in and out of this solution until the reagent solution has been substantially washed from the tip. Solution can be removed from the tip, after each dipping, by vacuum, compressed air spray, sponge, or the like.

The dispenser tip is now dipped in a second reagent well, and the filled tip is moved to a second selected array position in the first support. The process of dispensing reagent at each of the corresponding second-array positions is then carried as above. This process is repeated until an entire microarray of reagent solutions on each of the supports has been formed.

IV. Microarray Substrate

This section describes embodiments of a substrate having a microarray of biological polymers carried on the substrate surface. Subsection A describes a multi-cell substrate, each cell of which contains a microarray, and preferably an identical microarray, of distinct biopolymers, such as distinct polynucleotides, formed on a porous surface. Subsection B describes a microarray of distinct polynucleotides bound on a glass slide coated with a polycationic polymer.

A. Multi-Cell Substrate

Fig. 9 illustrates, in plan view, a substrate 110 constructed according to the invention. The substrate has an 8 x 12 rectangular array 112 of cells, such as cells 114, 116, formed on the substrate surface. With reference to Fig. 10, each cell, such as cell 114, in turn supports a microarray 118 of distinct biopolymers, such as polypeptides or polynucleotides at known, addressable regions of the microarray. Two such regions forming the microarray are indicated at 120, and correspond to regions, such as regions 42, forming the microarray of distinct biopolymers shown in Fig. 3.

The 96-cell array shown in Fig. 9 has typically array dimensions between about 12 and 244 mm in width and 8 and 400 mm in length, with the cells in the array having width and length dimension of 1/12 and 1/8 the array width and length dimensions, respectively, i.e., between about 1 and 20 in width and 1 and 50 mm in length.

The construction of substrate is shown cross-sectionally in Fig. 11, which is an enlarged sectional view taken along view line 124 in Fig. 9. The substrate includes a water-impermeable backing 126, such as a glass slide or rigid polymer sheet. Formed on the surface of the backing is a water-permeable film

128. The film is form d of a porous membrane material, such as nitrocellulose membrane, or a porous web material, such as a nylon, polypropylene, or PVDF porous polymer material. The thickness of the film is preferably between about 10 and 1000 μm. The film may be applied to the backing by spraying or coating uncured material on the backing, or by applying a preformed membrane to the backing. The backing and film may be obtained as a preformed unit from commercial source, *e.g.*, a plastic-backed nitrocellulose film available from Schleicher and Schuell Corporation.

With continued reference to Fig. 11, the film-covered surface in the substrate is partitioned into a desired array of cells by water-impermeable grid lines, such as lines 130, 132, which have infiltrated the film down to the level of the backing, and extend above the surface of the film as shown, typically a distance of 100 to 2000 μm above the film surface.

The grid lines are formed on the substrate by laying down an uncured or otherwise flowable resin or elastomer solution in an array grid, allowing the material to infiltrate the porous film down to the backing, then curing or otherwise hardening the grid lines to form the cell-array substrate.

One preferred material for the grid is a flowable silicone available from Loctite Corporation. The barrier material can be extruded through a narrow syringe (*e.g.*, 22 gauge) using air pressure or mechanical pressure. The syringe is moved relative to the solid support to print the barrier elements as a grid pattern. The extruded bead of silicone wicks into the pores of the solid support and cures to form a shallow waterproof barrier separating the regions of the solid support.

In alternativ embodiments, the barrier element can be a wax-based material or a thermoset material such as epoxy. The barrier material can also be a UV-curing polymer which is exposed to UV light after being printed onto the solid support. The barrier material may also be applied to the solid support using printing techniques such as silk-screen printing. The barrier material may also be a heat-seal stamping of the porous solid support which seals its pores and forms a water-impervious barrier element. The barrier material may also be a shallow grid which is laminated or otherwise adhered to the solid support.

In addition to plastic-backed nitrocellulose, the solid support can be virtually any porous membrane with or without a non-porous backing. Such membranes are readily available from numerous vendors and are made from nylon, PVDF, polysulfone and the like. In an alternative embodiment, the barrier element may also be used to adhere the porous membrane to a non-porous backing in addition to functioning as a barrier to prevent cross contamination of the assay reagents.

In an alternative embodiment, the solid support can be of a non-porous material. The barrier can be printed either before or after the microarray of biomolecules is printed on the solid support.

As can be appreciated, the cells formed by the grid lines and the underlying backing are water-impermeable, having side barriers projecting above the porous film in the cells. Thus, defined-volume samples can be placed in each well without risk of cross-contamination with sample material in adjacent cells. In Fig. 11, defined volumes samples, such as sample 134, are shown in the cells.

As noted above, each well contains a microarray of distinct biopolymers. In one general embodiment, th

microarrays in the well are identical arrays of distinct biopolymers, *e.g.*, different sequence polynucleotides. Such arrays can be formed in accordance with the methods described in Section II, by depositing a first selected polynucleotide at the same selected microarray position in each of the cells, then depositing a second polynucleotide at a different microarray position in each well, and so on until a complete, identical microarray is formed in each cell.

In a preferred embodiment, each microarray contains about 10^3 distinct polynucleotide or polypeptide biopolymers per surface area of less than about 1 cm^2. Also in a preferred embodiment, the biopolymers in each microarray region are present in a defined amount between about 0.1 femtomoles and 100 nanomoles. The ability to form high-density arrays of biopolymers, where each region is formed of a well-defined amount of deposited material, can be achieved in accordance with the microarray-forming method described in Section II.

Also in a preferred embodiments, the biopolymers are polynucleotides having lengths of at least about 50 bp, *i.e.*, substantially longer than oligonucleotides which can be formed in high-density arrays by schemes involving parallel, step-wise polymer synthesis on the array surface.

In the case of a polynucleotide array, in an assay procedure, a small volume of the labeled DNA probe mixture in a standard hybridization solution is loaded onto each cell. The solution will spread to cover the entire microarray and stop at the barrier elements. The solid support is then incubated in a humid chamber at the appropriate temperature as required by the assay.

Each assay may be conducted in an "open-face" format where no further sealing step is required, since the hybridization solution will be kept properly hydrated by the water vapor in the humid chamber. At the conclusion of the incubation step, the entire solid support containing the numerous microarrays is rinsed quickly enough to dilute the assay reagents so that no significant cross contamination occurs. The entire solid support is then reacted with detection reagents if needed and analyzed using standard colorimetric, radioactive or fluorescent detection means. All processing and detection steps are performed simultaneously to all of the microarrays on the solid support ensuring uniform assay conditions for all of the microarrays on the solid support.

B. Glass-Slide Polynucleotide Array

Fig. 5 shows a substrate 136 formed according to another aspect of the invention, and intended for use in detecting binding of labeled polynucleotides to one or more of a plurality distinct polynucleotides. The substrate includes a glass substrate 138 having formed on its surface, a coating of a polycationic polymer, preferably a cationic polypeptide, such as polylysine or polyarginine. Formed on the polycationic coating is a microarray 140 of distinct polynucleotides, each localized at known selected array regions, such as regions 142.

The slide is coated by placing a uniform-thickness film of a polycationic polymer, *e.g.*, poly-l-lysine, on the surface of a slide and drying the film to form a dried coating. The amount of polycationic polymer added is sufficient to form at least a monolayer of polymers on the glass surface. The polymer film is bound to surfac via electrostatic binding between

negative silyl-OH groups on the surface and charged amine groups in the polymers. Poly-l-lysine coated glass slides may be obtained commercially, e.g., from Sigma Chemical Co. (St. Louis, MO).

To form the microarray, defined volumes of distinct polynucleotides are deposited on the polymer-coated slide, as described in Section II. According to an important feature of the substrate, the deposited polynucleotides remain bound to the coated slide surface non-covalently when an aqueous DNA sample is applied to the substrate under conditions which allow hybridization of reporter-labeled polynucleotides in the sample to complementary-sequence (single-stranded) polynucleotides in the substrate array. The method is illustrated in Examples 1 and 2.

To illustrate this feature, a substrate of the type just described, but having an array of same-sequence polynucleotides, was mixed with fluorescent-labeled complementary DNA under hybridization conditions. After washing to remove non-hybridized material, the substrate was examined by low-power fluorescence microscopy. The array can be visualized by the relatively uniform labeling pattern of the array regions.

In a preferred embodiment, each microarray contains at least 10^3 distinct polynucleotide or polypeptide biopolymers per surface area of less than about 1 cm^2. In the embodiment shown in Fig. 5, the microarray contains 400 regions in an area of about 16 mm^2, or 2.5×10^3 regions/cm^2. Also in a preferred embodiment, the polynucleotides in the each microarray region are present in a defined amount between about 0.1 femtomoles and 100 nanomoles in the case of polynucleotides. As above, the ability to form high-

density arrays of this type, where each region is formed of a well-defined amount of deposited material, can be achieved in accordance with the microarray-forming method described in Section II.

Also in a preferred embodiments, the polynucleotides have lengths of at least about 50 bp, *i.e.*, substantially longer than oligonucleotides which can be formed in high-density arrays by various in situ synthesis schemes.

V. Utility

Microarrays of immobilized nucleic acid sequences prepared in accordance with the invention can be used for large scale hybridization assays in numerous genetic applications, including genetic and physical mapping of genomes, monitoring of gene expression, DNA sequencing, genetic diagnosis, genotyping of organisms, and distribution of DNA reagents to researchers.

For gene mapping, a gene or a cloned DNA fragment is hybridized to an ordered array of DNA fragments, and the identity of the DNA elements applied to the array is unambiguously established by the pixel or pattern of pixels of the array that are detected. One application of such arrays for creating a genetic map is described by Nelson, *et al.* (1993). In constructing physical maps of the genome, arrays of immobilized cloned DNA fragments are hybridized with other cloned DNA fragments to establish whether the cloned fragments in the probe mixture overlap and are therefore contiguous to the immobilized clones on the array. For example, Lehrach, *et al.*, describe such a process.

The arrays of immobilized DNA fragments may also be used for genetic diagnostics. To illustrate, an array containing multiple forms f a mutated gene or genes can be pr bed with a labeled mixture f a

patient's DNA which will preferentially interact with only one of the immobilized versions of the gene.

The detection of this interaction can lead to a medical diagnosis. Arrays of immobilized DNA fragments can also be used in DNA probe diagnostics. For example, the identity of a pathogenic microorganism can be established unambiguously by hybridizing a sample of the unknown pathogen's DNA to an array containing many types of known pathogenic DNA. A similar technique can also be used for unambiguous genotyping of any organism. Other molecules of genetic interest, such as cDNA's and RNA's can be immobilized on the array or alternately used as the labeled probe mixture that is applied to the array.

In one application, an array of cDNA clones representing genes is hybridized with total cDNA from an organism to monitor gene expression for research or diagnostic purposes. Labeling total cDNA from a normal cell with one color fluorophore and total cDNA from a diseased cell with another color fluorophore and simultaneously hybridizing the two cDNA samples to the same array of cDNA clones allows for differential gene expression to be measured as the ratio of the two fluorophore intensities. This two-color experiment can be used to monitor gene expression in different tissue types, disease states, response to drugs, or response to environmental factors. & An example of this approach is illustrated in Examples 2, described with respect to Fig. 8.

By way of example and without implying a limitation of scope, such a procedure could be used to simultaneously screen many patients against all known mutations in a disease gene. This invention could be used in the form of, for example, 96 identical 0.9 cm × 2.2 cm microarrays fabricated on a single 12 cm × 18 cm

sheet of plastic-backed nitrocellulose where each microarray could contain, for example, 100 DNA fragments representing all known mutations of a given gene. The region of interest from each of the DNA samples from 96 patients could be amplified, labeled, and hybridized to the 96 individual arrays with each assay performed in 100 microliters of hybridization solution. The approximately 1 thick silicone rubber barrier elements between individual arrays prevent cross contamination of the patient samples by sealing the pores of the nitrocellulose and by acting as a physical barrier between each microarray. The solid support containing all 96 microarrays assayed with the 96 patient samples is incubated, rinsed, detected and analyzed as a single sheet of material using standard radioactive, fluorescent, or colorimetric detection means (Maniatas, *et al.*, 1989). Previously, such a procedure would involve the handling, processing and tracking of 96 separate membranes in 96 separate sealed chambers. By processing all 96 arrays as a single sheet of material, significant time and cost savings are possible.

The assay format can be reversed where the patient or organism's DNA is immobilized as the array elements and each array is hybridized with a different mutated allele or genetic marker. The gridded solid support can also be used for parallel non-DNA ELISA assays. Furthermore, the invention allows for the use of all standard detection methods without the need to remove the shallow barrier elements to carry out the detection step.

In addition to the genetic applications listed above, arrays of whole cells, peptides, enzymes, antibodies, antigens, receptors, ligands, phospholipids, polymers, drug cogener preparations or

chemical substances can be fabricated by the means described in this invention for large scale screening assays in medical diagnostics, drug discovery, molecular biology, immunology and toxicology.

The multi-cell substrate aspect of the invention allows for the rapid and convenient screening of many DNA probes against many ordered arrays of DNA fragments. This eliminates the need to handle and detect many individual arrays for performing mass screenings for genetic research and diagnostic applications. Numerous microarrays can be fabricated on the same solid support and each microarray reacted with a different DNA probe while the solid support is processed as a single sheet of material.

The following examples illustrate, but in no way are intended to limit, the present invention.

Example 1

Genomic-Complexity Hybridization to Micro DNA Arrays Representing the Yeast *Saccharomyces cerevisiae* Genome with Two-Color Fluorescent Detection

The array elements were randomly amplified PCR (Bohlander, *et al.*, 1992) products using physically mapped lambda clones of *S. cerevisiae* genomic DNA templates (Riles, *et al.*, 1993). The PCR was performed directly on the lambda phage lysates resulting in an amplification of both the 35 kb lambda vector and the 5-15 kb yeast insert sequences in the form of a uniform distribution of PCR product between 250-1500 base pairs in length. The PCR product was purified using Sephadex G50 gel filtration (Pharmacia, Piscataway, NJ) and concentrated by evaporation to dryness at room temperature overnight. Each of the 864 amplified

lambda clones was rehydrated in 15 μl of 3 × SSC in preparation for spotting onto the glass.

The micro arrays were fabricated on microscope slides which were coated with a layer of poly-l-lysine (Sigma). The automated apparatus described in Section IV loaded 1 μl of the concentrated lambda clone PCR product in 3 × SSC directly from 96 well storage plates into the open capillary printing element and deposited ~5 nl of sample per slide at 380 micron spacing between spots, on each of 40 slides. The process was repeated for all 864 samples and 8 control spots. After the spotting operation was complete, the slides were rehydrated in a humid chamber for 2 hours, baked in a dry 80° vacuum oven for 2 hours, rinsed to remove un-absorbed DNA and then treated with succinic anhydride to reduce non-specific adsorption of the labeled hybridization probe to the poly-l-lysine coated glass surface. Immediately prior to use, the immobilized DNA on the array was denatured in distilled water at 90° for 2 minutes.

For the pooled chromosome experiment, the 16 chromosomes of *Saccharomyces cerevisiae* were separated in a CHEF agarose gel apparatus (Biorad, Richmond, CA). The six largest chromosomes were isolated in one gel slice and the smallest 10 chromosomes in a second gel slice. The DNA was recovered using a gel extraction kit (Qiagen, Chatsworth, CA). The two chromosome pools were randomly amplified in a manner similar to that used for the target lambda clones. Following amplification, 5 micrograms of each of the amplified chromosome pools were separately random-primer labeled using Klenow polymerase (Amersham, Arlington Heights, IL) with a lissamine conjugated nucleotide analog (Dupont NEN, Bost n, MA) for the pool containing the six largest chromosomes, and with a fluorescein

conjugated nucleotide analog (BMB) for the pool containing smallest ten chromosomes. The two pools were mixed and concentrated using an ultrafiltration device (Amicon, Danvers, MA).

Five micrograms of the hybridization probe consisting of both chromosome pools in 7.5 μl of TE was denatured in a boiling water bath and then snap cooled on ice. 2.5 μl of concentrated hybridization solution (5 × SSC and 0.1% SDS) was added and all 10 μl transferred to the array surface, covered with a cover slip, placed in a custom-built single-slide humidity chamber and incubated at 60° for 12 hours. The slides were then rinsed at room temperature in 0.1 × SSC and 0.1%SDS for 5 minutes, cover slipped and scanned.

A custom built laser fluorescent scanner was used to detect the two-color hybridization signals from the 1.8 × 1.8 cm array at 20 micron resolution. The scanned image was gridded and analyzed using custom image analysis software. After correcting for optical crosstalk between the fluorophores due to their overlapping emission spectra, the red and green hybridization values for each clone on the array were correlated to the known physical map position of the clone resulting in a computer-generated color karyotype of the yeast genome.

Figure 6 shows the hybridization pattern of the two chromosome pools. A red signal indicates that the lambda clone on the array surface contains a cloned genomic DNA segment from one of the largest six yeast chromosomes. A green signal indicates that the lambda clone insert comes from one of the smallest ten yeast chromosomes. Orange signals indicate repetitive sequences which cross hybridized to both chromosome pools. Control spots on the array confirm that the hybridization is specific and reproducible.

The physical map locations of the genomic DNA fragments contained in each of the clones used as array elements have been previously determined by Olson and co-workers (Riles, *et al.*) allowing for the automatic generation of the color karyotype shown in Figure 7. The color of a chromosomal section on the karyotype corresponds to the color of the array element containing the clone from that section. The black regions of the karyotype represent false negative dark spots on the array (10%) or regions of the genome not covered by the Olson clone library (90%). Note that the largest six chromosomes are mainly red while the smallest ten chromosomes are mainly green matching the original CHEF gel isolation of the hybridization probe. Areas of the red chromosomes containing green spots and vice-versa are probably due to spurious sample tracking errors in the formation of the original library and in the amplification and spotting procedures.

The yeast genome arrays have also been probed with individual clones or pools of clones that are fluorescently labeled for physical mapping purposes. The hybridization signals of these clones to the array were translated into a position on the physical map of yeast.

Example 2

Total cDNA Hybridized to Micro Arrays of cDNA Clones with Two-Color Fluorescent Detection

24 clones containing cDNA inserts from the plant *Arabidopsis* were amplified using PCR. Salt was added to the purified PCR products to a final concentration of 3 × SSC. The cDNA clones were spotted on poly-l-lysine coated microscope slides in a manner similar to Exampl 1. Among the cDNA clones was a clone

representing a transcription factor HAT 4, which had previously been used to create a transgenic line of the plant *Arabidopsis*, in which this gene is present at ten times the level found in wild-type *Arabidopsis* (Schena, *et al.*, 1992).

Total poly-A mRNA from wild type *Arabidopsis* was isolated using standard methods (Maniatis, *et al.*, 1989) and reverse transcribed into total cDNA, using fluorescein nucleotide analog to label the cDNA product (green fluorescence). A similar procedure was performed with the transgenic line of *Arabidopsis* where the transcription factor HAT4 was inserted into the genome using standard gene transfer protocols. cDNA copies of mRNA from the transgenic plant are labeled with a lissamine nucleotide analog (red fluorescence). Two micrograms of the cDNA products from each type of plant were pooled together and hybridized to the cDNA clone array in a 10 microliter hybridization reaction in a manner similar to Example 1. Rinsing and detection of hybridization was also performed in a manner similar to Example 1. Fig. 8 show the resulting hybridization pattern of the array.

Genes equally expressed in wild type and the transgenic *Arabidopsis* appeared yellow due to equal contributions of the green and red fluorescence to the final signal. The dots are different intensities of yellow indicating various levels of gene expression. The cDNA clone representing the transcription factor HAT4, expressed in the transgenic line of *Arabidopsis* but not detectably expressed in wild type *Arabidopsis*, appears as a red dot (with the arrow pointing to it), indicating the preferential expression of the transcription factor in the red-labeled transgenic *Arabidopsis* and the r lative lack of expression of the

transcription factor in the gr en-labeled wild type *Arabidopsis*.

An advantage of the microarray hybridization format for gene expression studies is the high partial concentration of each cDNA species achievable in the 10 microliter hybridization reaction. This high partial concentration allows for detection of rare transcripts without the need for PCR amplification of the hybridization probe which may bias the true genetic representation of each discrete cDNA species.

Gene expression studies such as these can be used for genomics research to discover which genes are expressed in which cell types, disease states, development states or environmental conditions. Gene expression studies can also be used for diagnosis of disease by empirically correlating gene expression patterns to disease states.

Example 3

Multiplexed Colorimetric Hybridization on a Gridded Solid Support

A sheet of plastic-backed nitrocellulose was gridded with barrier elements made from silicone rubber according to the description in Section IV-A. The sheet was soaked in 10 × SSC and allowed to dry. As shown in Fig. 12, 192 M13 clones each with a different yeast inserts were arrayed 400 microns apart in four quadrants of the solid support using the automated device described in Section III. The bottom left quadrant served as a negative control for hybridization while each of the other three quadrants was hybridized simultaneously with a different oligonucleotide using the open-face hybridization technology described in Section IV-A. The first two and last four elements of

each array are positive controls for the colorimetric detection step.

The oligonucleotides were labeled with fluorescein which was detected using an anti-fluorescein antibody conjugated to alkaline phosphatase that precipitated an NBT/BCIP dye on the solid support (Amersham). Perfect matches between the labeled oligos and the M13 clones resulted in dark spots visible to the naked eye and detected using an optical scanner (HP ScanJet II) attached to a personal computer. The hybridization patterns are different in every quadrant indicating that each oligo found several unique M13 clones from among the 192 with a perfect sequence match. Note that the open capillary printing tip leaves detectable dimples on the nitrocellulose which can be used to automatically align and analyze the images.

Although the invention has been described with respect to specific embodiments and methods, it will be clear that various changes and modification may be made without departing from the invention.

IT IS CLAIMED:

1. A method of forming a microarray of analyte-assay regions on a solid support, where each region in the array has a known amount of a selected, analyte-specific reagent, said method comprising,

(a) loading a solution of a selected analyte-specific reagent in a reagent-dispensing device having an elongate capillary channel (i) formed by spaced-apart, coextensive elongate members, (ii) adapted to hold a quantity of the reagent solution and (iii) having a tip region at which aqueous solution in the channel forms a meniscus,

(b) tapping the tip of the dispensing device against a solid support at a defined position on the surface, with an impulse effective to break the meniscus in the capillary channel and deposit a selected volume of solution on the surface, and

(c) repeating steps (a) and (b) until said array is formed.

2. The method of claim 1, wherein said tapping is carried out with an impulse effective to deposit a selected volume in the volume range between 0.01 to 100 nl.

3. The method of claim 1, wherein said channel is formed by a pair of spaced-apart tapered elements.

4. The method of claim 1, for forming a plurality of such arrays, wherein step (b) is applied to a selected position on each of a plurality of solid supports at each repeat cycle proceeding step (c).

5. The method of claim 1, which further includes, after performing steps (a) and (b) at least one time, reloading the reagent-dispensing device with a new reagent solution by the steps of (i) dipping the capillary channel of the device in a wash solution, (ii) removing wash solution drawn into the capillary channel, and (iii) dipping the capillary channel into the new reagent solution.

6. Automated apparatus for forming a microarray of analyte-assay regions on a plurality of solid supports, where each region in the array has a known amount of a selected, analyte-specific reagent, said apparatus comprising

(a) a holder for holding, at known positions, a plurality of planar supports,

(b) a reagent dispensing device having an open capillary channel (i) formed by spaced-apart, coextensive elongate members (ii) adapted to hold a quantity of the reagent solution and (iii) having a tip region at which aqueous solution in the channel forms a meniscus,

(c) positioning means for positioning the dispensing device at a selected array position with respect to a support in said holder,

(d) dispensing means for moving the device into tapping engagement against a support with a selected impulse, when the device is positioned at a defined array position with respect to that support, with an impulse effective to break the meniscus of liquid in the capillary channel and deposit a selected volume of solution on the surface, and

(e) control means for controlling said positioning and dispensing means.

7. The apparatus of claim 6, wh rein said dispensing means is effective to move said dispensing device against a support with an impulse effective to deposit a selected volume in the volume range between 0.01 to 100 nl.

8. The apparatus of claim 6, wherein said channel is formed by a pair of spaced-apart tapered elements.

9. The apparatus of claim 6, wherein the control means operates to (i) place the dispensing device at a loading station, (ii) move the capillary channel in the device into a selected reagent at the loading station, to load the dispensing device with the reagent, and (iii) dispense the reagent at a defined array position on each of the supports on said holder.

10. The apparatus of claim 6, wherein the control device further operates, at the end of a dispensing cycle, to wash the dispensing device by (i) placing the dispensing device at a washing station, (ii) moving the capillary channel in the device into a wash fluid, to load the dispensing device with the fluid, and (iii) remove the wash fluid prior to loading the dispensing device with a fresh selected reagent.

11. The apparatus of claim 6, wherein said device is one of a plurality of such devices which are carried on the arm for dispensing different analyte assay reagents at selected spaced array positions.

12. A substrate with a surface having a microarray of at least 10^3 distinct polynucleotide or polypeptide biopolymers per 1 cm^2 surface area, each

distinct biopolymer sample (i) being disposed at a separate, defined position in said array, (ii) having a length of at least 50 subunits, and (iii) being present in a defined amount between about 0.1 femtomole and 100 nanomoles.

13. The substrate of claim 12, wherein said surface is glass slide coated with polylysine, and said biopolymers are polynucleotides.

14. The substrate of claim 12, wherein said substrate has a water-impermeable backing, a water-permeable film formed on the backing, and a grid formed on the film, where said grid (i) is composed of intersecting water-impervious grid elements extending from said backing to positions raised above the surface of said film, and (ii) partitions the film into a plurality of water-impervious cells, where each cell contains such a biopolymer array.

15. A substrate with a surface array of sample-receiving cells, comprising

a water-impermeable backing,

a water-permeable film formed on the backing, and a grid formed on the film, said grid being composed of intersecting water-impervious grid elements extending from said backing to positions raised above the surface of said film.

16. The substrate of claim 15, wherein the cells of the array each contain an array of biopolymers.

17. A substrate for use in detecting binding of labeled biopolymers to one or more of a plurality distinct polynucleotides, comprising

a non-porous, glass substrate,

a coating of a cationic polymer on said substrate, and

an array of distinct polynucleotides to said coating, where each biopolymer is disposed at a separate, defined position in a surface array of biopolymers.

18. A method of detecting differential expression of each of a plurality of genes in a first cell type with respect to expression of the same genes in a second cell types, said method comprising

producing fluorescence-labeled cDNA's from mRNA's isolated from the two cells types, where the cDNA's from the first and second cells are labeled with first and second different fluorescent reporters,

adding a mixture of the labeled cDNA's from the two cell types to an array of polynucleotides representing a plurality of known genes derived from the two cell types, under conditions that result in hybridization of the cDNA's to complementary-sequence polynucleotides in the array; and

examining the array by fluorescence under fluorescence excitation conditions in which (i) polynucleotides in the array that are hybridized predominantly to cDNA's derived from one of the first and second cell types give a distinct first or second fluorescence emission color, respectively, and (ii) polynucleotides in the array that are hybridized to substantially equal numbers of cDNA's derived from the first and second cell types give a distinct combined fluorescence emission color, respectively,

wherein the relative expression of known genes in the two cell types can b determined by the observed fluorescence emission color of each spot.

19. The method of claim 18, wherein the array of polynucleotides is formed on a substrate with a surface having an array of at least 10^2 distinct polynucleotide or polypeptide biopolymers in a surface area of less than about 1 cm^2, each distinct biopolymer (i) being disposed at a separate, defined position in said array, (ii) having a length of at least 50 subunits, and (iii) being present in a defined amount between about .1 femtomole and 100 nmoles.

20. The method of claim 19, wherein said surface is a glass slide coated with polylysine, and said biopolymers are polynucleotides non-covalently bound to said polylysine.

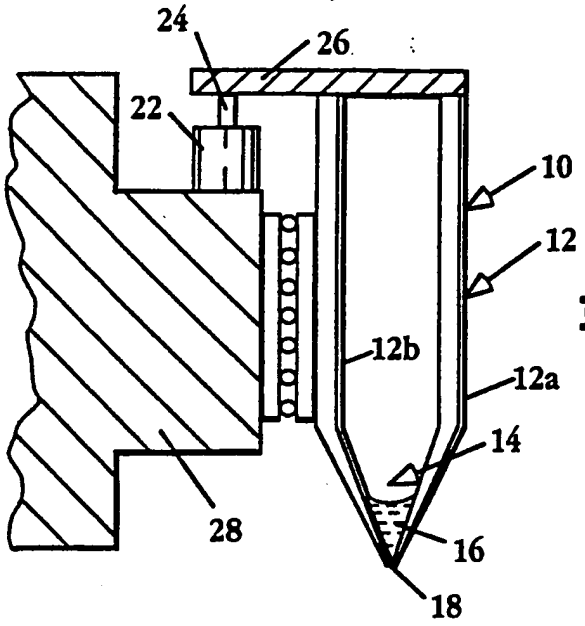


Fig. 1

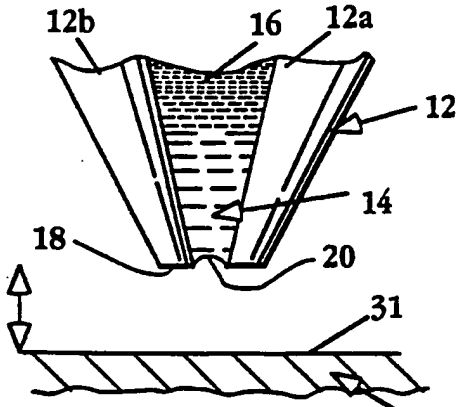


Fig. 2A

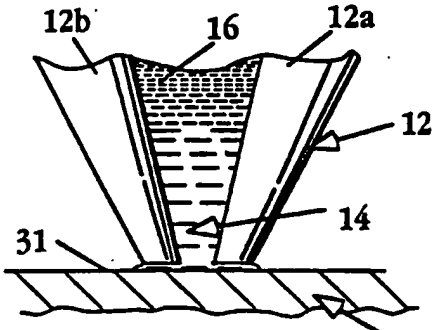


Fig. 2B

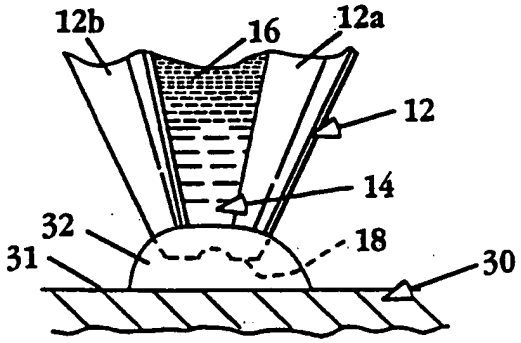


Fig. 2C

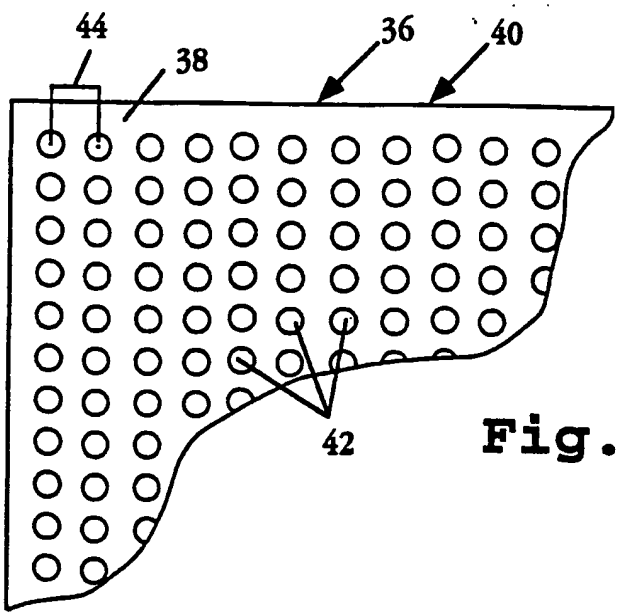

44
38
36
40
42

Fig. 3

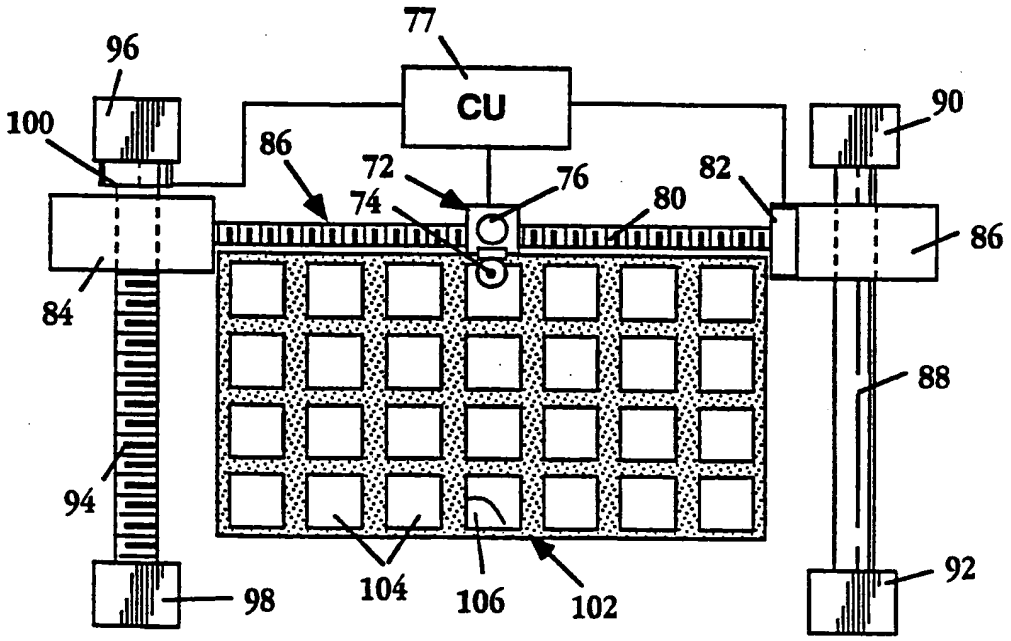

77
CU
96
100
86
72
74
76
80
82
90
86
84
88
94
98
104
106
102
92

Fig. 4

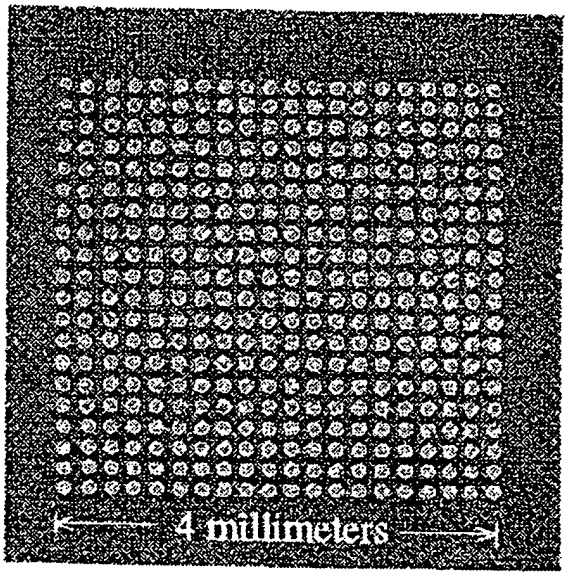


Fig. 5

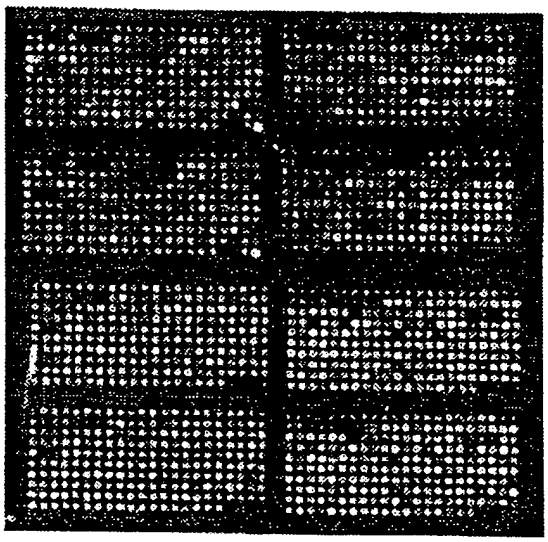


Fig. 6

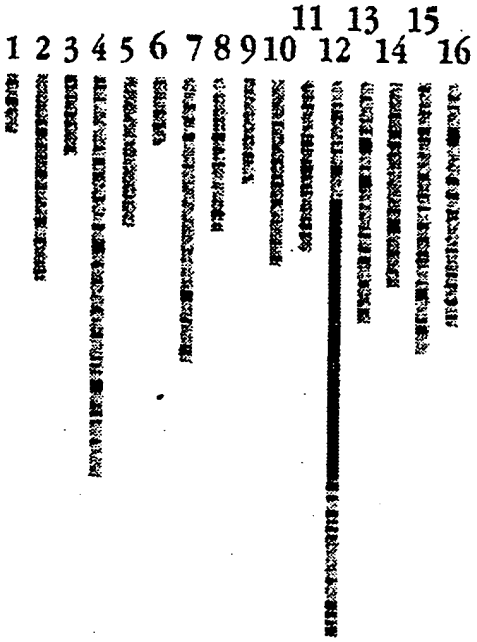


Fig. 7

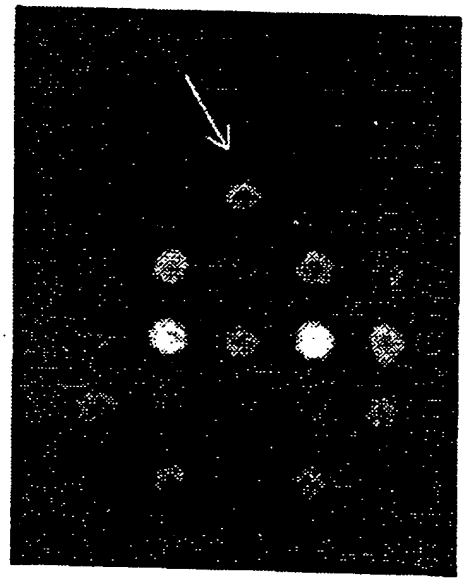


Fig. 8

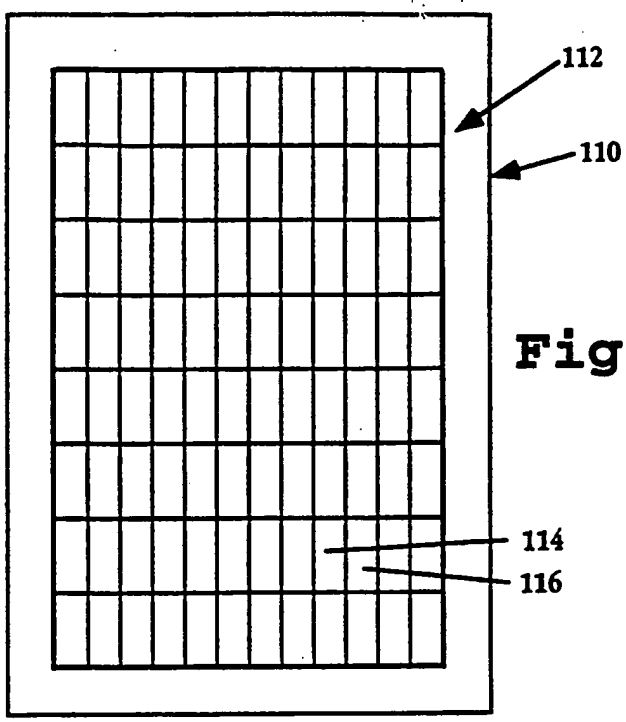


Fig. 9

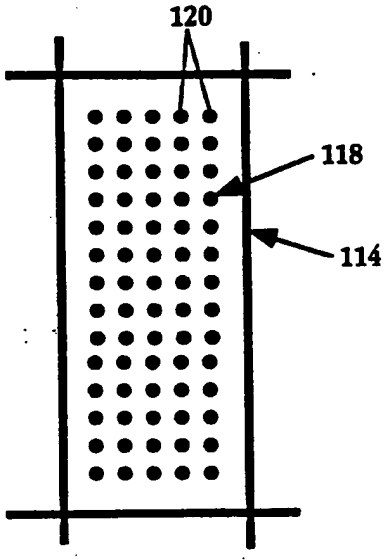


Fig. 10

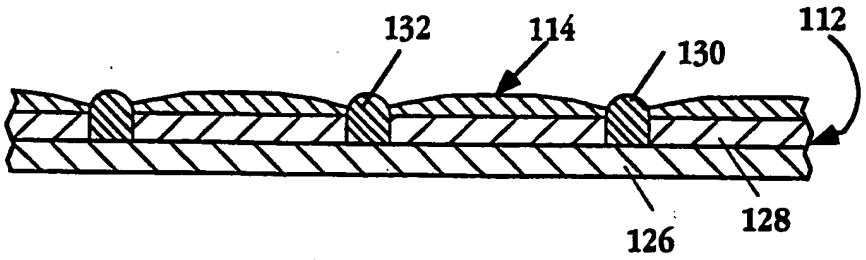


Fig. 11

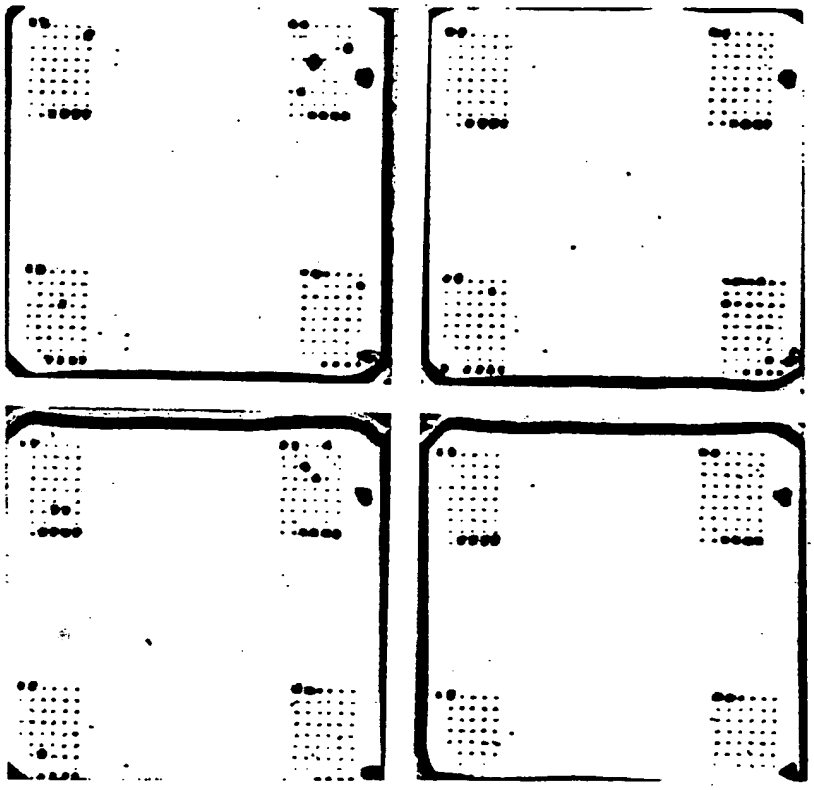


Fig. 12

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US95/07659

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) :G01N 33/543, 33/68
US CL :435/6; 436/518
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 422/57; 435/4,6,973; 436/518,524,527,531,805,809

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A,P	US, A, 5,338,688 (DEEG ET AL) 16 August 1994, see entire document	1-17
A	US, A, 5,204,268 (MATSUMOTO) 20 April 1993, see entire document.	6-11
A	US, A, 4,071,315 (CHATEAU) 31 January 1978, see entire document.	12-17
A	US, A, 5,100,777 (CHANG) 31 March 1992, see entire document.	12-17
A	US, A, 5,200,312 (OPRANDY) 06 April 1993, see entire document.	12-17

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier document published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search 15 SEPTEMBER 1995	Date of mailing of the international search report 06 OCT 1995
Name and mailing address of the ISA/US Commissioner of Patents and Trademarks Box PCT Washington, D.C. 20231 Facsimile No. (703) 305-3230	Authorized officer CHRISTOPHER CHIN Telephone No. (703) 308-0196

Form PCT/ISA/210 (second sheet)(July 1992)*

nature genetics

volume 14 no. 4 december 1996

DNA chips, diagnostics and genomics

Rieger syndrome

QTLs and epistasis

A BRCA1-binding protein

Nature Genetics

Editor
Kevin Davies

Assistant Editors
Laurie Goodman
Bette Phimister

Production Editor
J. Stuart Griffith

Assistant Production Editor
Ken Krattenmaker

Editorial Assistant
Janelle Bolden

Washington Bureau Chief
Barbara J. Culliton

Editorial Office
545 National Press Building
Washington DC 20045
Tel: (202) 626-2513
Fax: (202) 626-0970
email: natgen naturedc.com

WWW: genetics.nature.com

Cover art: Ken Krattenmaker

Nature Genetics (ISSN 1061-4036) is published monthly by Nature Publishing Co., headquartered at 345 Park Avenue South, New York, NY 10010, which is owned by Nature America Inc., a subsidiary of Macmillan Magazines Ltd. of London. Editorial Office: Nature Genetics, 545 National Press Building, Washington DC 20045. Telephone (202) 626 2513, Fax (202) 626 0970, e-mail natgen@naturedc.com. North American Advertising: Nature Genetics, 345 Park Avenue South, New York, NY 10010. Telephone (212) 726 9200, Fax (212) 696 9606. European Advertising: Nature Genetics, Porters South, Crinan Street, London N1 9XW, UK. Telephone (0171) 833 4000, Fax (0171) 843 4596. New subscriptions, renewals, changes of address, back issues and all customer service questions should be addressed to: Nature Genetics Subscription Dept, 345 Park Avenue South, New York, NY 10010, USA. Telephone (212) 726 9200, Fax (212) 696 9606. Outside North America: Nature Genetics, Macmillan Magazines Ltd, Porters South, Crinan Street, London N1 9XW, UK. Telephone (0171) 833 4000, Fax (0171) 843 4596. Annual subscription rates: US/Canada: US$495 (institutional/corporate), US$195 (individual making personal payment), Canada add 7% for GST, BN: 14091 1595 RT; UK: £350 (institutional/corporate): £175 (individual making personal payment). Japan: Contact Nature Japan, K.K., Shin-Mitsuke Bldg. 4F, 3-6 Ichigaya Tamachi, Shinjuku-ku, Tokyo 162. Reprints: Nature Genetics Reprints Dept, 345 Park Avenue South, New York, NY 10010, USA. Telephone (212) 726 9200, Fax (212) 696 9606. Periodicals postage paid at New York, NY and at additional mailing offices. POSTMASTER: Send address changes to Nature Genetics Subscription Dept, 345 Park Avenue South, New York, NY 10010. Executive Officers of Nature America, Inc.: Nicholas Byam Shaw, Chairman of the Board; Mary Waltham, President; Edward Vals, Secretary-Treasurer. Published in Japan by Nature Japan KK, Shin-Mitsuke Building, 3F, 3-6 Ichigaya Tamachi, Shinjuku-ku, Tokyo 162, Japan. Printed in the USA by Cadmus Journal Services.

Nature Publishing Co.
345 Park Avenue South
10th floor
New York, NY 10010-1707
Tel: (212) 726-9200
Fax: (212) 696-9606

President–Publisher
Mary Waltham

Vice President Sales
Manon Delaney

Vice President Marketing
James A. Skowrenski

American Advertising Sales Manager
Sande T. Giaccone (New York)

European Advertising Sales Manager
Kathryn Wayman (London)

Classified Advertising Sales Manager
Erika A. Simon (New York)
Mike Grant (London)

Assistant Classified Sales Manager
Benjamin Crowe (New York)

Production & Information Systems Director
Nick Kemp

Circulation Manager
Moira Musto (New York)
Nic Harman (London)

Group Marketing Manager
Gina Dzurenda

Macmillan Magazines Ltd
Porters South
Crinan Street
London N1 9XW
Tel: 44 (0)171 833 4000
Fax: 44 (0)171 843 4596

Managing Director
Kay Barker

Publishing Director
Andy Sutherland

Editor-in-Chief,
Nature publications
Philip Campbell

Art Director
Jane Walker

Nature Japan KK
Shin-Mitsuke Bldg
3-6 Ichigaya Tamachi
Shinjuku-ku
Tokyo 162
Telephone 03 3267 8751
Fax 03 3267 8746

Publisher
David Swinbanks

Use of a cDNA microarray to analyse gene expression patterns in human canc r

Joseph DeRisi[1*], Lolita Penland[2] & Patrick O. Brown[2] (Group 1); Michael L. Bittner[3*], Paul S. Meltzer[3], Michael Ray[3], Yidong Chen[3], Yan A. Su[3] & Jeffrey M. Trent[3] (Group 2)

The development and progression of cancer[1–3] and the experimental reversal of tumorigenicity[4,5] are accompanied by complex changes in patterns of gene expression. Microarrays of cDNA provide a powerful tool for studying these complex phenomena[6–8]. The tumorigenic properties of a human melanoma cell line, UACC-903, can be suppressed by introduction of a normal human chromosome 6, resulting in a reduction of growth rate, restoration of contact inhibition, and suppression of both soft agar clonogenicity and tumorigenicity in nude mice[4,5,9]. We used a high density microarray of 1,161 DNA elements to search for differences in gene expression associated with tumour suppression in this system. Fluorescent probes for hybridization were derived from two sources of cellular mRNA [UACC-903 and UACC-903(+6)] which were labelled with different fluors to provide a direct and internally controlled comparison of the mRNA levels corresponding to each arrayed gene. The fluorescence signals representing hybridization to each arrayed gene were analysed to determine the relative abundance in the two samples of mRNAs corresponding to each gene. Previously unrecognized alterations in the expression of specific genes provide leads for further investigation of the genetic basis of the tumorigenic phenotype of these cells.

DNA microarrays, containing 1,161 total elements, including 870 different cDNAs and controls[9–11] (see Methods), were printed robotically onto a glass microscope slide in four quadrants covering an area of about 1 cm^2 (Fig. 1). We prepared fluorescent cDNA probes using total poly (A)$^+$ mRNA from UACC-903 cells and UACC-903(+6) cells by labelling with a green and red fluor, respectively. A mixture of the two flourescently labelled probes was hybridized to the DNA microarray. This comparative hybridization method, coupled with the doping of synthetic standards and an estimation of statistically significant deviation for local background variance allowed a direct and quantitative comparison of the relative abundance of individual DNA sequences in this complex sample[6–8]. We added a set of synthetic poly (A)$^+$-tailed 'mRNAs' to the purified mRNA from each cell line as internal standards to assist in quantitation and estimation f experimental variation introduced during labelling and reading. Targets complementary to these standards were included, in duplicate, on the microarray. Based on these standards, mRNA species comprising 1:10,000 of the mass of the poly (A)$^+$ RNA could readily be detected.

In a representative two-colour fluorescent scan of all 1,161 [illegible] (Fig. 2a), the green spots correspond t genes preferentially expressed in the tumorigenic UACC-903 cell line, and the reddish spots correspond to genes preferentially expressed in the non-tumorigenic UACC-903(+6) cell line. Genes expressed at approximately equal levels in the two cell lines appear yellow or brown. A portion of the array at higher magnification highlights the diverse pattern of differential expression observed (Fig. 2*b*). In Fig. 2*c*, rectangles corresponding to specific array elements are coloured to reproduce the hue and intensity of the fluorescent signal at each element. The hybridization signals from a duplicated set of genes are shown juxtaposed, to illustrate the reproducibility of the hybridization signals for each gene.

To address the possibility that an apparent difference in expression might result from experimental variables unrelated to the difference in chromosomal composition between the two cell lines, we examined the variance in expression for 90 'housekeeping' genes. We selected these genes based on the assumption that they would not be differentially expressed between the two cell lines. The averaged red/green ratio for this subset of genes was 1.13. The averaged red/green ratio for the set of five internal standards was 0.97 (n = 10). The variability in the expression level of the housekeeping genes probably overestimates the experimental variability in measuring differential expression. As a conservative standard, an absolute fluorescent signal (red or green) with an intensity greater than that observed at the control array elements containing total human genomic DNA was considered to represent specific hybridization. Gene-specific hybridization was therefore only considered significantly different between samples if the following two criteria were met: i) the signal intensity (green or red) exceeded this threshold; and ii) the logarithm of the red/green fluorescence signal ratio differed by ≥3 S.D. from the mean logarithm of this ratio for the 'housekeeping' gene panel (that is, ratios <0.52 or >2.4).

By these criteria, mRNA levels for 15/870 (1.7%) genes were significantly diminished, while the mRNA levels for 63/870 (7.3%) genes were significantly increased in association with suppression of tumorigenicity by introduction of chromosome 6. To test the reliability of microarray hybridization results in identifying differentially expressed genes, we analysed 16 genes by northern analysis. In each case, the results of northern analysis corroborated the differential gene expression identified by microarray hybridization (Fig. 3).

Significant differences in expression between these two cell lines identified several genes as candidates for determining features of the tumorigenic phenotype of the melanoma cells. For example, among the genes detected with significantly higher expression (>10-fold) in the tumorigenic cells was the human brown locus protein (TRP1/melanoma antigen gp75). This is the most abundant glycoprotein in melanocytic cells and a critical melanosome membrane protein[12,13]. Additionally, its expression is reduced when melanoma cell lines are induced to differentiate by treatment with HMBA[12,13]. Also expressed at a significantly higher level was a spliced variant of the mRNA encoding myelin PLP/DM20. This is widely expressed in neural crest derived cells in early development and has been suggested to play a role in cell–cell signaling during development[14].

[illegible] lev-

[1]*Howard Hughes Medical Institute,* [2]*Department of Biochemistry, Stanford University Medical Center, Stanford, California 94305, USA* [3]*Laboratory of Cancer Genetics, National Center for Human Genome Research, National Institutes of Health, Bethesda, Maryland 20892, USA*

**J.D. & M.L.B. contributed equally to this work.*

Correspondence should be addressed to P.B. or J.T.
e-mail: pbrown@cmgm.stanford.edu jtrent@nchgr.
nih.gov

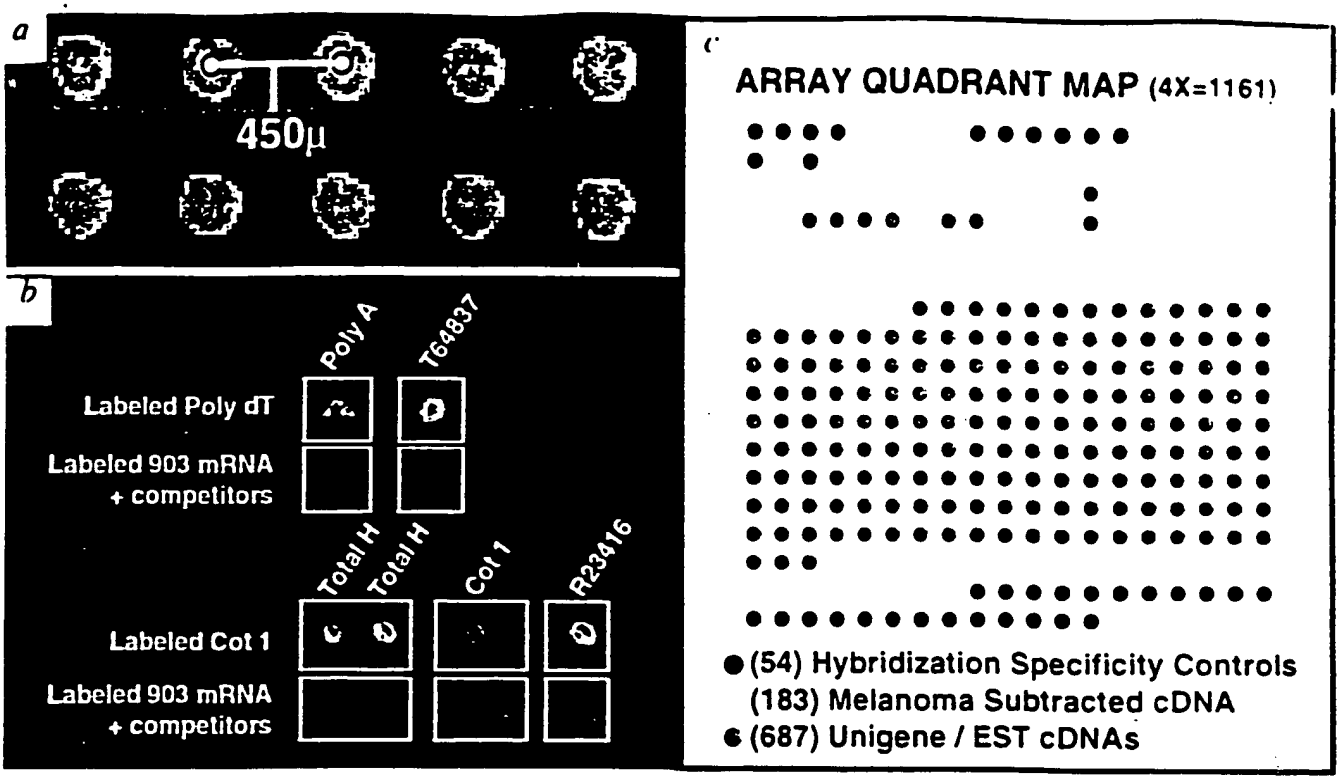


Fig. 1 Properties of cDNA microarrays. a, A fluorescent scan of DNA printed onto a poly-lysine coated slide. The DNA is stained with a DNA-specific fluorescent dye, YOYO. The center-to-center spacing of adjacent spots is 450 μ, allowing the potential for up to 10,000 spots/2.54 X 7.62 cm microscope slide. b, Efficient blocking of hybridization to DNA repeats. Hybridization of fluorescein-labelled poly (dT)* to arrays in the absence of competitor produces strong hybridization to immobilized poly (dA)* as well as to some cDNAs, such as the EST T64827 shown. Rhodamine-labelled cDNA (red) from the UACC-903 cell line hybridized in the presence of poly (dA)* blocker shows little if any signal at either site (Total H = total human). Similarly, hybridization with fluorescein-labelled Cot1 DNA in the absence of competitor produces bright signal on immobilized Cot1 DNA, total human DNA and at some cDNA elements (presumed to contain highly repeated sequences, such as R23416); while Rhodamine-labelled cDNA (red) from the UACC-903 cell line produces little if any signal at these locations when hybridized in the presence of excess unlabelled poly (dA)*, and human Cot1 DNA. The absence of signal at some cDNA locations following UACC-903 cDNA hybridizations also indicates that the PCR-amplified, plasmid vector sequences at all cDNA targets do not contribute significant hybridization signal. c, Schematic of the array organisation. Robotic printing from 96 well microtiter trays was carried out with 4 print heads, spaced to fit into 4 adjacent microtiter wells. This maps the contents of each tray into four separate quadrants on the glass slide. A colour-coded map of the general distribution of target types in each of the resulting quadrants is shown.

els were elevated by the addition of a normal chromosome 6 (17 genes) are known to be activated by IFN-γ, a cardinal proinflammatory cytokine that, among other activities, induces expression of the gene products of the MHC class II locus. For example, the mRNA encoding monocyte chemotactic protein 1 (MCAF/MCP1), a cytokine that induces monocyte chemotaxis and activation[15,16], was more than 10-fold less abundant in the tumorigenic cell line. In the skin, MPC1 is critical in the regulation of cutaneous monocyte trafficking[16–17], and elevated expression plays a role in suppression of tumour growth and metastasis[19–21]. The mechanism by which these interferon-γ regulated genes are induced in UACC-903 cells by transfer of a normal chromosome 6 remains to be determined. It is worth noting, however, that the interferon-γ receptor gene is localized to the distal long arm of human chromosome 6.

Finally, several genes that showed >10-fold higher expression in the suppressed UACC-903(+6) cells have previously been recognized in other models of tumour suppression. Most notably, there was elevated expression of the mRNA encoding WAF1 (p21), a key mediator of tumour suppression by p53 (ref. 18). The p21 protein had previously been identified as a melanoma differentiation-associated antigen (termed mda-6)[19,20]. In melanoma cell lines suppressed for metastasis by the introduction of chromosome 6, expression of WAF1 (p21) mRNA and protein correlates inversely with metastatic potential[20].

These results provide a wide view of the diverse systems that are altered in this model system of tumorigenicity, and focus attention on specific gene products and pathways that may be of particular importance in this tumour type.

Our ability to classify human cancers in a way that reflects the underlying molecular pathology or that anticipates their potential for progression or response to treatment, remains primitive. Using cDNA microarrays to define alterations in gene expression associated with a specific cancer may be an efficient way to uncover clues to the specific molecular derangements that contribute to its pathogenesis and thus identify potential targets for therapeutic intervention. Moreover, recognition of pathognomonic alterations in gene expression might provide a basis for improved diagnosis and molecular classification of cancers and thus allow selection of the most appropriate therapeutic strategies.

Public databases of human expressed gene sequences contain partial sequences of at least 40,000 different human genes[11], and efforts to develop a human transcript map have developed rapidly[21]. Based on the high yield of information obtained using an array of <1,000 different genes, a more comprehensive survey of gene expression patterns, using a more complete array of human genes, will likely provide a rich source of new and useful insights into human biology and a deeper understanding of the gene pathways involved in the pathogenesis of cancer and other diseases.

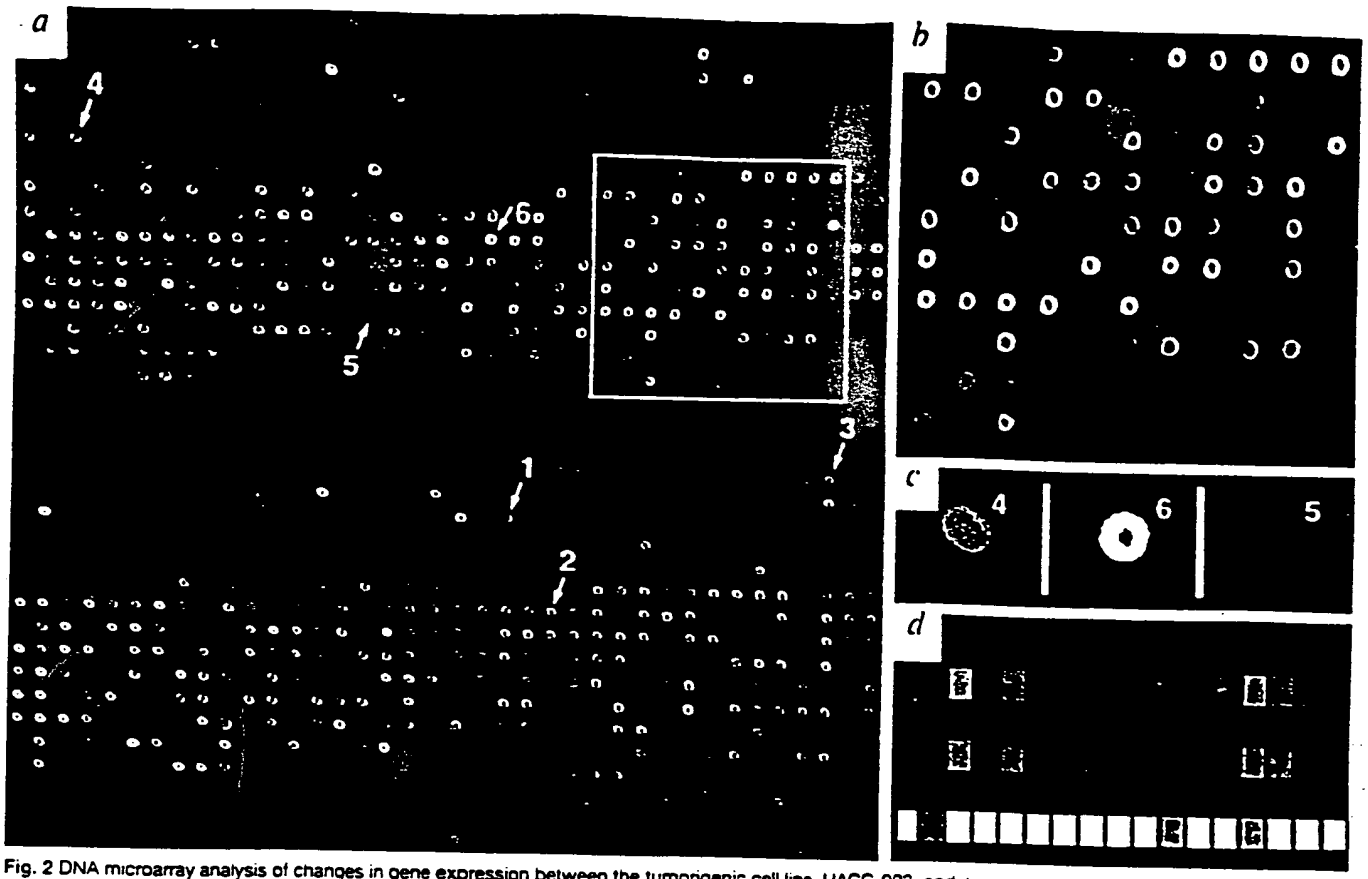


Fig. 2 DNA microarray analysis of changes in gene expression between the tumorigenic cell line, UACC-903, and its non-tumorigenic derivative, UACC903(+6), derived by introduction of a normal chromosome 6. *a*, A ratio image of the results of simultaneous hybridization of Rhodamine110-labelled cDNA (green) from UACC-903 and Cy3-labelled cDNA (orange-red) from UACC-903(+6) to a microarray. To produce this image, the scan images corresponding to each fluorescent probe were combined as the appropriate colour channels in a single image. Arrows indicate the location within the array of the corresponding genes analysed by northern blotting (Fig. 3). *b*, A magnified image of the area of the array boxed in white in (*a*). *c*, Magnified image of three cDNAs identified by arrows in (*a*), representing the cDNAs for: left, *MCAF/MCP-1* (r/g ratio >10); centre, *β-actin* (r/g ratio 1.04); and right, *α-1-antichymotrypsin* (r/g ratio 0.2) [see Fig. 3]. *d*, simplified representation of ratio hybridization results. Quantitative fluorescence intensity data is extracted from each array target. The average target colour ratio determines the hue of each box and the average intensity determines the brightness of each box. In this image, the order of the boxes corresponds to their original order in the microtiter plate from which they were printed. Duplicate printings of the same plate can be examined side by side, as in the first two rows shown here, to assess reproducibility of the hybridization results (see text). Numbered arrows indicate the location within the array corresponding to genes analysed by northern blotting in Fig. 3.

Methods

Generation of microarrays, hybridization, scanning. The preparation of coated microscope slides and subsequent robotic printing of DNA was carried out in a manner similar to that described[7]. Briefly, pre-cleaned glass slides were treated with poly-L-lysine solution (Sigma) to form an adhesive surface for printing. PCR products, purified by ethanol purification, were resuspended in 3x SSC. A custom built arraying robot picked up and deposited small volumes (~5 nanoliters) of DNA onto the slides. After printing, the slides were washed in a 0.2% SDS solution. The remaining bound DNA was denatured by submerging the slides in 95 °C distilled water for 2 min followed by a brief wash with 95% ethanol. DNA was UV crosslinked to the slides (Stratagene Stratalinker, 60 mJ). To prevent non-specific probe binding, the slides were blocked by rinsing in a solution of 70 mM succinic anhydride dissolved in 0.1 M boric acid pH 8.0, containing 35% 1-methyl-2-pyrrolidinone (Aldrich). Additional protocols and parts list pertaining to microarray fabrication can be obtained from http://cmgm.stanford.edu/pbrown.

Purified, labelled cDNA was resuspended in 11 μl of 3.5× SSC containing 4 μg of poly (dA)$^+$ DNA, 2.5 μg *E. coli* tRNA, 4 μg of human Cot1 DNA (Gibco BRL), and 0.3 μl of 10% SDS. Prior to hybridization, the solution was boiled for 2 min then allowed to cool to room temperature. Hybridization was carried out at 62 °C for ~14 h in a water bath. Prior to scanning, slides were washed in 2× SSC, 0.2% SDS for 5 min and 0.2× SSC for 1 min.

[illegible] focal laser microscope built by S. Smith with software written by N. Ziv. A separate scan, using the appropriate excitation line, was done for each of the two fluorophores used. Data was collected at a maximum resolution of 9 microns/pixel with 12 bits of depth.

Probe preparation and labelling. RNA was extracted from cells using the Triazol reagent (LTI Inc.), following the manufacturer's directions. cDNA probes were synthesized from singly oligo dT-selected (Pharmacia) mRNA pools. Fluorescently labelled cDNA was prepared from mRNA by oligo dT-primed polymerization using SuperScript II reverse transcriptase (LTI Inc.). The pool of nucleotides in the labelling reaction was 0.5 mM dGTP, dATP and dCTP and 0.2 mM dTTP. Fluorescent nucleotides, Rhodamine 110 dUTP (Perkin Elmer Cetus) or Cy3 dUTP (Amersham), were present at 0.1 mM. Probes were purified by gel chromatography (BioSpin 6/BioRad) and ethanol precipitation.

Selection of cDNA elements and generation of control templates. Synthetic cDNAs were prepared by cloning random *Bam*HI and *Hin*dIII ended fragments of *E. coli* DNA in the vector pSP64 poly (A)$^+$ (Promega), linearizing isolated plasmid DNA with *Eco*RI and synthesizing poly (A)$^+$ tailed RNA complementary to the insert from the resident SP6 promoter (Promega). Prior to use, the synthesized RNAs were selected on oligo dT cellulose. The largest group of cDNAs consisted of 674 cDNA clones from the UNIB arrayed normalized infant brain [illegible]. These clones were selected to [illegible] every NIB library member that corresponded to a named gene according

Fig. 3 Northern hybridization substantiating the consistency of the cDNA microarray results. Corresponding locations within the cDNA microarray illustrated in Fig. 2*a* are provided for 1) *Waf-1/p21*; 2) *MARCKS*; 3) *collagenase*; 4) *MCAF/MCP-1*; 5) *α-1-antichymotrypsin*; and 6) *β-actin*. The signal detected by a radio-labelled *β-actin* probe represents a control for loading variance, with a red/green ratio observed on the cDNA microarray (Fig. 2*a,c*) for *β-actin* of 1.04.

to the UniGene EST clustering system[21,22]. The second largest group of clones consisted of 183 sequenced cDNA clones generated by subtraction of cDNA from the chromosome-6 suppressed non-tumorigenic UACC-903 (+6) cell line with cDNA from its parental tumorigenic cell line UACC-903 (ref. 9). Approximately 100 additional genes (total 870 genes arrayed) were obtained from EST libraries on the basis of their expression pattern (tissue specific, and so on). Each array included the following hybridization controls: plasmid vector, lambda, ΦX174 phage, total human DNA, human Cot1 DNA, and poly (A)$^+$. The synthetic standards used for normalization of signals in each wavelength were also arrayed. Controls were included in each quadrant of the array to assess the reproducibility of the hybridization signal. Two plates of cDNA clones (derived from the UACC-903 subtracted library) were also arrayed in duplicate. Fidelity of the Unigene array relative to dbEST was tested by sequencing of a random sample of 11 clones used for microarray construction. All sequences were identical with the corresponding dbEST entries. Additionally, each microarrayed cDNA from the UACC-903 subtracted library was sequenced. A listing of cDNAs comprising this microarray which were derived from the Unigene and 'housekeeping' panel can be obtained from http://www.nih.gov/DIR/LCG/ARRAY/expn.html.

Northern blot analysis. Total RNA, 10 μg per lane, was electrophoresed in 1.2% agarose-formaldehyde gels and transferred onto nylon membrane (Hybond-N$^+$, Amersham) by capillary blotting overnight. For DNA probes insert fragments from the Soares IMAGE cDNA library[10] were obtained by vector PCR for p21, MARCKS, α-1-antichymotrypsin and β-actin. Probes for fibroblast collagenase and MCAF/MCP-1 were isolated from a UACC-903(+6) enriched cDNA library[9] with all probes labelled by random priming. Filters were washed to a stringency of 0.1× SSC at 42 °C for 20 min.

Web sites. http://cmgm.stanford.edu/pbrown for protocols and parts list pertaining to microarray fabrication. http://www.nchgr.nih.gov/DIR/LCG/ARRAY/expn.html for a listing of cDNAs comprising this microarray which were derived from the Unigene and 'housekeeping' panel.

Acknowledgements

Work in P.O.B.'s laboratory is supported in part by the Howard Hughes Medical Institute and National Center for Human Genome Research (HG00450). We would like to acknowledge the excellent technical and graphic assistance of X. He, T. Hofmann, Y. Jiang, J. Leuders, D. Leja and B. Walker. J.D. was supported by NIH grant 2T32BM07276-21. P.O.B. is an assistant investigator of the Howard Hughes Medical Institute.

Received 15 October; accepted 8 November, 1996.

1. Vogelstein, B. & Kinzler, K.W. The multistep nature of cancer. *Trends Genet.* 9, 138–141 (1993).
2. Weinberg, R.A. The molecular basis of oncogenes and tumor suppressor genes. *Ann. NY Acad. Sci.* 758, 331–338 (1995).
3. Levine, A.J. The tumor suppressor genes. *Annu. Rev. Biochem.* 62, 623–651 (1993).
4. Trent, J.M. *et al.* Tumorigenicity in human melanoma cell lines controlled by introduction of human chromosome 6. *Science* 247, 568–571 (1990).
5. Su, Y. *et al.* Reversion of monochromosome-mediated suppression of tumorigenicity in malignant melanoma by retroviral transduction. *Cancer Res.* 56, 3186–3191 (1996).
6. Schena, M., Shalon, D., Davis, R.W., & Brown, P.O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467–470 (1995).
7. Shalon, D., Smith, S.J. & Brown, P.O. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res.* 6, 639–645 (1996).
8. Schena, M. *et al.* Parallel human genome analysis: microarray-based expression of 1000 genes. *Proc. Natl. Acad. Sci. USA* 93, 10539–11286 (1996).
9. Ray, M.E., Su, Y.A., Meltzer, P.S. & Trent, J.M. Isolation and characterization of genes associated with chromosome-6 mediated tumor suppression in human malignant melanoma. *Oncogene* 12, 2527–2533 (1996).
10. Soares, M.B. *et al.* Construction and characterization of a normalized cDNA library. *Proc. Natl. Acad. Sci. USA* 91, 9228–9232 (1994).
11. Boguski, M.S. & Schuler, G.D. ESTablishing a human transcript map *Nature Genet.* 10, 369–371 (1995).
12. Vijayasaradhi, S., Doskoch, P.M., Wolchok, J. & Houghton, A.N. Melanocyte differentiation marker gp75, the brown locus protein, can be regulated independently of tyrosinase and pigmentation. *J. Invest. Dermatol.* 105, 113–119 (1995).
13. Vijayasaradhi, S., Xu, Y., Bouchard, B. & Houghton, A.N. Intracellular sorting and targeting of melanosomal membrane proteins: identification of signals for sorting of the human brown locus protein, gp 75. *J. Invest. Dermatol.* 130, 807–820 (1995).
14. Nakao, J. *et al.* Expression of proteolipid protein gene is directly associated with secretion of a factor influencing oligodendrocyte development. *J. Neurochem.* 6455, 2396–2403 (1995).
15. Graves, D.T., Barnhill, R., Galanopoulos, T. & Antoniades, H.N. Expression of monocyte chemotactic protein-1 in human melanoma in vivo. *Am. J. Pathol.* 140, 9–14 (1992).
16. Kristensen, M S., Deleuran, B.W., Larsen, C.G., Thestrup-Pedersen, K. & Paludan, K. Expression of monocyte chemotactic and activating factor (MCAF) in skin related cells. A comparative study. *Cytokine* 5, 520–524 (1993).
17. Huang, S., Xie, K. Singh, R K., Gutman, M. & Bar-Eli, M. Suppression of tumor growth and metastasis of murine renal adenocarcinoma by syngeneic fibroblasts genetically engineered to secrete the JE/MCP-1 cytokine. *J. Interferon Cytokine Res.* 15, 655–665 (1995).
18. El-Deiry, W.S. *et al.* WAF1, a potential mediator of p53 tumor suppression. *Cell* 75, 817–825 (1993).
19. Miele, M.E. *et al.* Metastasis suppressed, but tumorigenicity and local invasiveness unaffected, in the human melanoma cell line MeWuSo after introduction of human chromosomes 1 or 6. *Mol. Carcinog.* 15, 284–299 (1996).
20. Jiang, H. *et al.* The melanoma differentiation-associated gene mda-6, which encodes the cyclin-dependent kinase inhibitor p21, is differentially expressed during growth, differentiation and progression in human melanoma cells. *Oncogene* 10, 1855–1864 (1995).
21. Schuler, G.D. *et al.* A gene map of the human genome. *Science* 274, 540–546 (1996).
22. Lennon, G., Auffray, C., Polymeropoulos, M. & Soares, M.B. The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* 33, 151–152 (1996).

1088-9051

GENOME RESEARCH

July 1996 Volume 6 Number 7

INCLUDIN PCR

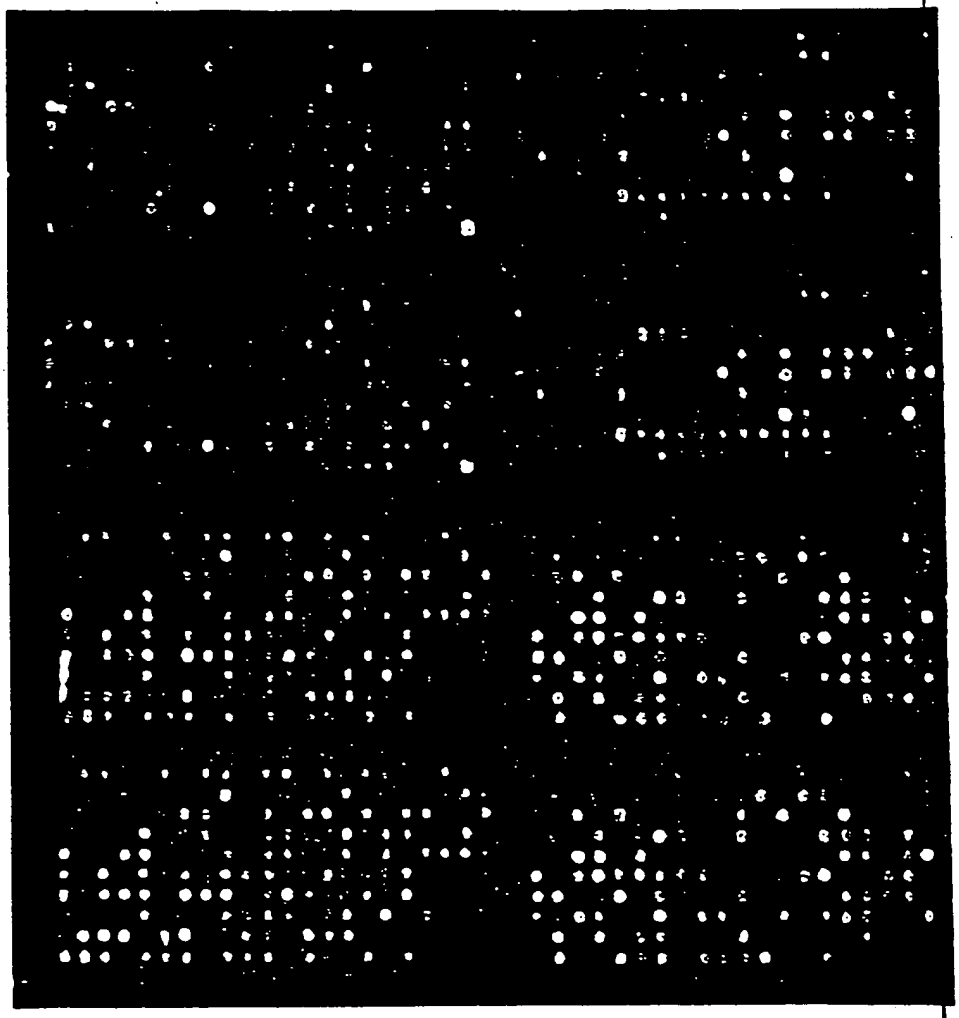
IBD Mapping in Livestock

Sequence of 500-kb *Rhizobium* Replicon

Human Y Chromosome Haplotypes

BAC Mapping of Extrachromosomal Structure

DNA Microarray System



Cold Spring Harbor Laboratory Press

Editorial office: Cold Spring Harbor Laboratory Press, 1 Bungtown Road, Cold Spring Harbor, New York 11724-2203. Phone 516-367-8492; FAX 516-367-8334.

GENOME RESEARCH (ISSN 1054-9803) is published m nthly for $495 (U.S. institutional; $545 R.O.W.), $95 (individual making personal payment; $145 R.O.W., includes airlift) by Cold Spring Harbor Laboratory Press, 1 Bungtown Road, Cold Spring Harbor, New York 11724. Periodicals class postage pending is paid at Cold Spring Harbor and additional mailing offices. POSTMASTER: Send address changes to C ld Spring Harbor Laboratory, 10 Skyline Drive, Plainview, New York 11803-2500.

Subscriptions: Barbara Terry, Subscription Manager. Personal: U.S. $95; R.O.W. $145 (includes airlift). Institutional: U.S. $495; R.O.W. $545 (includes airlift). Orders may be sent to Cold Spring Harbor Laboratory Press, Fulfillment Department, 10 Skyline Drive, Plainview, New York 11803-2500. Telephone: Continental U.S. and Canada 1-800-843-4388; all other locations 516-349-1930. FAX 516-349-1946. Personal subscriptions must be prepaid by personal check, credit card, or money order. Claims for missing issues must be received within 4 m nths f issue date.

Advertising: Teresa Tiganis, Advertising Manager, C ld Spring Harbor Laborat ry Press, 1 Bungtown Road, C ld Spring Harb r, New Y rk 11724-2203. Ph ne: 516-367-8351; FAX 516-367-8334.

GENOME RESEARCH

Volume 6 Number 7
July 1996

RESEARCH PAPERS

LETTERS

(continued)

GENOME METHODS

COVER DNA microarrays for analyzing complex DNA samples. Shown is a two-color fluorescent scan of an 1.8-cm × 1.8-cm yeast array of λ clones of yeast genomic DNA. (For details, see Shalon et al., p. 639.)

A DNA Microarray System for Analyzing Complex DNA Samples Using Two-color Fluorescent Probe Hybridization

Dari Shalon,[1,4] Stephen J. Smith,[3] and Patrick O. Brown[1,2,5]

[1]Howard Hughes Medical Institute and Departments of [2]Biochemistry and [3]Molecular and Cellular Physiology, Stanford University, Stanford, California 94305

Detecting and determining the relative abundance of diverse individual sequences in complex DNA samples is a recurring experimental challenge in analyzing genomes. We describe a general experimental approach to this problem, using microscopic arrays of DNA fragments on glass substrates for differential hybridization analysis of fluorescently labeled DNA samples. To test the system, 864 physically mapped λ clones of yeast genomic DNA, together representing >75% of the yeast genome, were arranged into 1.8-cm × 1.8-cm arrays, each containing a total of 1744 elements. The microarrays were characterized by simultaneous hybridization of two different sets of isolated yeast chromosomes labeled with two different fluorophores. A laser fluorescent scanner was used to detect the hybridization signals from the two fluorophores. The results demonstrate the utility of DNA microarrays in the analysis of complex DNA samples. This system should find numerous applications in genome-wide genetic mapping, physical mapping, and gene expression studies.

Many problems in genome analysis depend on determining what specific sequences are represented in a complex DNA or RNA sample and at what abundance, for example, what genes are represented in a specific chromosome band or YAC clone, what intervals are amplified or deleted in a particular cancer cell, or what genes are expressed in specific cells under specific conditions. As a general approach to this problem, we have developed a system for making microarrays of DNA samples on glass substrates, probing them by hybridization with complex fluorescent-labeled probes, and using a laser-scanning microscope to detect the fluorescent signals representing hybridization. Fluorescent labeling allows for simultaneous hybridization and separate detection of the hybridization signal from two or more probes. This in turn allows very accurate and reliable measurement of the relative abundance of specific sequences in two complex samples.

RESULTS

Array Hybridization Pattern

Figure 1 shows the two-color fluorescent scan of a yeast genomic array following hybridization with a mixed probe consisting of lissamine-labeled DNA from the 6 largest yeast chromosomes together with fluorescein-labeled DNA from the 10 smallest yeast chromosomes. A red color indicates that yeast sequences present in the lissamine-labeled hybridization probe hybridized to an array element. A yellow-green color indicates that yeast sequences present in the fluorescein-labeled hybridization probe hybridized to an array element. An orange color indicates cross-hybridization of both chromosome pools to an array element (e.g., dispersed repetitive elements, such as Ty1 elements).

Each clone was spotted twice, resulting in duplicate hybridization patterns in adjacent quadrants of the array. Control DNA spots, which were randomly amplified in the same manner as the λ clone array elements, are located in the bottom corner of each quadrant. "A" points to a pair of spots containing total yeast genomic DNA. These spots appear orange because both chromosome pools hybridized to yeast genomic DNA. The negative controls are as follows: "B" points to a pair of spots of wild-type λ DNA, "C" points to a pair of human genomic DNA spots, and "D" points to a pair of φX174 DNA spots. The lack of a hybridization signal at these three negative control spots indicates that the hybridization was specific for yeast sequences.

[4]Present address: Synteni, Inc., Palo Alto, California 94305.
[5]Corresponding author.
E-MAIL pbrown cmgm.stanford.edu; http://cmgm.stanford.edu/pbrown; FAX (415) 723-1399.

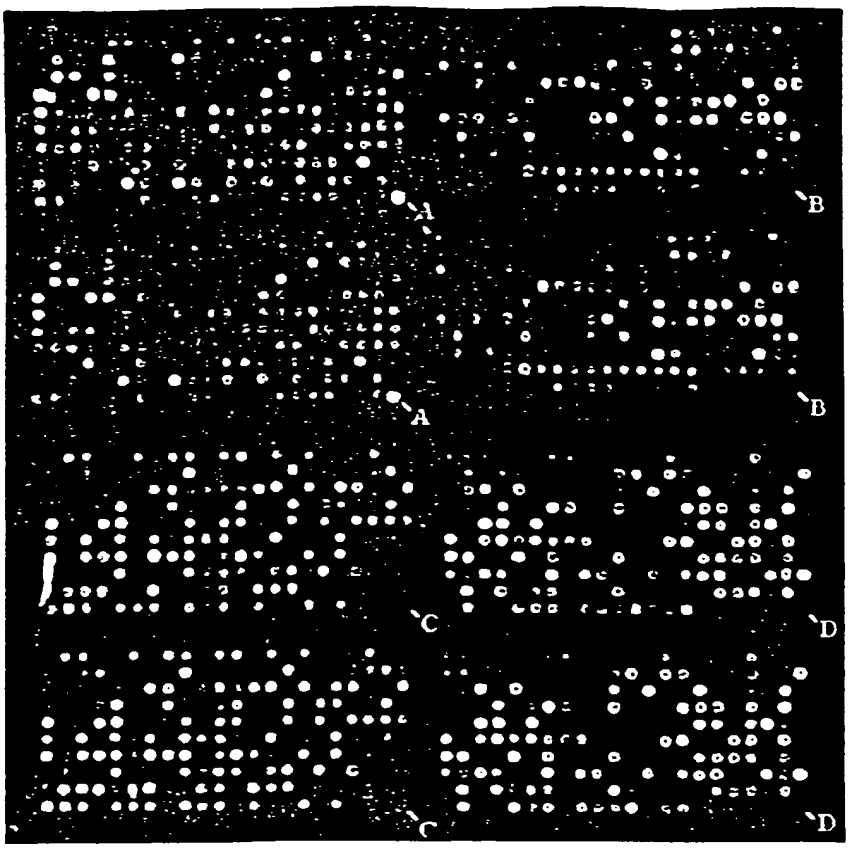


Figure 1 Two-color fluorescent scan of a 1.8-cm × 1.8-cm yeast array of λ clones of yeast genomic DNA. The DNA spots are spaced at a distance of 380 μm from center to center. A probe mixture consisting of DNA fr m the 6 largest yeast chromosomes (4, 7, 12, 13, 15, 16) labeled with lissamine (red dots) and DNA from the 10 smallest yeast chromosomes (1, 2, 3, 5, 6, 8, 9, 10, 11, 14) labeled with fluorescein (yellow-green dots) was hybridized to the array. A pair of yeast genomic DNA spots (*A*) served as a positive control. The three negative controls are λ DNA (*B*), human genomic DNA (*C*), and ϕX174 DNA (*D*).

Karyotype Depiction of the Array Hybridization Pattern

The inserts contained in the arrayed λ clones have been mapped physically (Riles et al. 1993). The clones are arrayed in a random but known order on the array. Therefore, using the identity of each clone along with its physical map information, the pattern of hybridization to the yeast array can be represented in the form of a karyotype of the yeast genome, as shown in Figure 2. The color of any segment of the ideogram representing an individual chromosome on the karyotype is directly determined by the ratio of red and green hybridization signals at the array positions of the corresponding clones. The lengths of the discrete colored segments of each chromosome correspond to the physical lengths of the yeast inserts. The chromosome segments colored black represent either intervals of the genome that are not represented by clones in the library (90%) or false-negative hybridization signals on the array (10%). Most of these false negatives are attributable to failures of the PCR amplification of the λ clones, though occasional failures of the arraying process or nonuniform surface preparation could account for a small fraction of the false-negative signals. The large gap on chromosome 12 is the region coding for ribosomal DNA that was not represented among the arrayed clones. Genomic intervals represented by overlapping clones were assigned a color based on the hybridization signals of only one of the overlapping clones, chosen at random.

Note that in this representation of a yeast karyotype, the largest six chromosomes are mainly colored red. This indicates that most of the arrayed clones that were mapped previously to these six large chromosomes hybridized primarily to the lissamine-labeled probe prepared from the corresponding purified chromosomes. Conversely, the smallest 10 chromosomes are mainly colored green in this image, matching the original CHEF gel isolation of the chromosomes used as the hybridization probe. The experiment was repeated with the yeast genome split into six discrete chromosome pools containing 2–4 chromosomes per pool using CHEF gel electrophoresis. The chromosomes in each pool were extracted from the gel, amplified, and fluorescently labeled. The six chromosome pools were hybridized to six separate yeast arrays. Forty-four λ clones gave a positive hybridization signal on all six arrays indicating that they contain yeast repetitive sequences (data not shown). These 44 clones and 10 clones with very weak hybridization signals were not included in the data set used to produce this karyotype.

There were ~40 anomalous clones, which appear in this karyotype representation as green bands on the otherwise red chromosomes or red

Figure 2 Computer-generated ideogram representing a karyotype of *S. cerevisiae,* based on the normalized hybridization signals from the array shown in Fig. 1. Note that the 6 largest chromosomes are mainly red and the 10 smallest chromosomes are mainly green. Black stripes represent intervals not represented by clones in the array or for which the corresponding clones gave false-negative hybridization signals.

bands on the otherwise green chromosomes. Four randomly chosen examples of these anomalous clones were analyzed by hybridizing the clones to vertical strips cut from a Southern blot of CHEF gel-separated yeast chromosomes. In each case, the hybridization patterns of the anomalous clones corroborated the chromosomal locations assigned by the microarray hybridization results (data not shown). Two clones that were thought to map to the 10 smallest chromosomes were found to hybridize preferentially to the probe representing the 6 largest chromosomes and thus appear as anomalous red bands on the karyotype. Both hybridized to one of the six largest chromosomes on the Southern blot. Similarly, two clones that appear as anomalous green bands on the karyotype were found to hybridize to one of the 10 smallest chromosomes on the Southern blot. Thus, the anomalous clones are probably the result of sample tracking errors or, possibly, of errors in the published restriction-digest-based physical map on which the karyotype representation was based (Riles et al. 1993).

DISCUSSION

The DNA microarray hybridization system reported here is conceptually and functionally similar to fluorescent in situ hybridization (FISH) to metaphase chromosomes, with three important differences. First, the target elements of the microarrays can, in principle, be any length or composition, from megabase YAC clones or microdissected chromosome bands to individual cDNA clones, to short oligonucleotides. This versatility allows the user to choose characteristics, such as the mapping resolution and genetic complexity of each array element, to suit a particular application. Second, the hybridization signals are localized to discrete elements of known size and location, making them easier to identify and quantitate than the hybridization signals from irregularly shaped metaphase spreads. Third, microarrays are more consistent and potentially amenable to automated production, hybridization, and data analysis than metaphase spreads.

Arrays of DNA samples on porous membranes, for example, dot blots, have long been used as a basic tool in molecular biology. Dot-blot membranes are usually at least 8 × 12 cm in size, require the use of milliliter volumes of hybridization solution, and are limited, owing to autofluorescence and scattering, to radioactive, chemiluminescent, and colorimetric hybridization detection methods (Ross et al. 1992). Microarrays made on glass surfaces, on the other hand, can be mass-produced and are comparatively inexpensive, convenient, and compatible with fluorescent hybridization detection methods. Furthermore, a glass surface, when appropriately treated, has very low nonspecific binding of labeled hybridization probes, resulting in lower backgrounds than are encountered typically with porous membranes. For hybridizations with very complex probes, the concentration of the labeled probe DNA is a limiting factor in the sensitivity of the assay. Minimizing the volume of the probe solution in a hybridization, by restricting the target to a small area and by using a nonporous substrate, makes it practical to achieve very high probe concentrations.

One important advantage of fluorescently labeled probes is that, unlike most radioactive and chemiluminescent signals, fluorescent signals do not disperse and therefore allow for very dense array spacing. A unique, and probably the m st important, advantage of fluorescent probes is that the hybridization signals from two or more differently labeled probes hybridized to the same target element can be detected separately. In this way, two-color hybridization detection allows for a direct and quantitative comparison f the

abundance of specific sequences between two probe mixtures that are hybridized competitively to a single array. The absolute intensity of a hybridization signal at a particular element in an array can vary owing to experimental factors such as variations in the amount of DNA deposited on the array, variations in the hybridization or wash conditions between experiments, or variations in the hybridization characteristics of the different DNA sequences on the array. The ratio of the two signals at any element in an array, however, is relatively insensitive to these confounding factors because they affect both probe mixtures equivalently. This ratio therefore accurately reflects the relative abundance of the cognate sequence in the two probe samples. This is the principle underlying the technique of comparative genomic hybridization (CGH), which is used to detect changes in the copy number of specific chromosomes or chromosomal regions (Kallioniemi et al. 1992). CGH is based on measuring the relative fluorescent hybridization intensities of two genomic-complexity hybridization probes, for example, probes representing genomic DNA from normal and affected tissue samples, which are labeled with two distinct fluorophores and hybridized simultaneously to a metaphase spread. DNA microarray representations of the human genome may provide a more convenient and higher resolution alternative to metaphase chromosomes for CGH.

Cross-hybridization between related sequences is an important problem faced by any hybridization-based assay, including the DNA microarray assay described here. Studies are now in progress to quantitate the extent of cross-hybridization between related sequences of varying homology and length, in DNA microarray hybridizations. The stringency of hybridization and washing can be controlled by varying the salt concentration and temperature as in conventional membrane-based hybridizations. Cross-hybridization caused by repetitive sequences can be minimized by prehybridization of the probe or array with vast excess of unlabeled copies of the repetitive sequences.

Alternative methods have been described for making microarrays of very short DNA sequences, involving photolithography (Pease et al. 1994) or physical masking (Maskos and Southern 1992) methods. These in situ synthesis methods are inherently limited to low complexity array elements consisting of oligonucleotides. For complex-probe hybridizations, the specificity of hybridization is improved by using DNA fragments substantially longer than oligonucleotides. Moreover, the in situ synthesis approaches to array fabrication depend on pri r knowledge of the sequence to be recognized by each array element. The approach described here makes microarrays by transferring tiny volumes of DNA samples from microwell storage plates to a solid substrate. Thus, nucleic acids (or other molecules) of virtually any length or any origin can be arrayed, and knowledge of their sequences is not required.

The arrays used in these experiments do not represent the maximal achievable density of elements. We have found that the spacing between the spots can be decreased by shrinking the contact area of the printing tip and by increasing the hydrophobicity of the glass surface. Microarrays with 100-μm feature size have been tested successfully in pilot experiments (data not shown). Assuming the projected availability of the appropriate physically mapped human genomic clones (Hudson et al. 1995), arrays at 100-μm spacing would allow for 10,000 discrete intervals of the human genome to be represented in a 1-cm^2 array. Such an array could be used for mapping at a resolution of <0.5 Mb. Experiments are in progress to explore the feasibility of such arrays.

Our initial motivation for developing these microarrays arose from the need for abundant and inexpensive genomic arrays for genomic mismatch scanning (GMS) (Nelson et al. 1993), a method of genetic linkage analysis based on identification of the regions of "identity by descent" between affected relative pairs using a single complex-probe hybridization to an array of genomic clones. Experiments using these arrays to map quantitative trait loci in yeast by GMS are currently in progress (J. deRisi, D. Lashkari, L. Penland, L. McAllister, J. McCusker, R. Davis, and P.O. Brown, unpubl.).

Microarrays of cDNA clones, prepared using the system described here, have been used for quantitative monitoring of gene expression patterns in *Arabidopsis* (Schena et al. 1995), *S. cerevisiae* (D. Lashkari, J. deRisi, L. Penland, P.O. Brown, and R. Davis, unpubl.), and human tissues (J. deRisi, M. Bittner, P. Meltzer, L. Penland, J. Trent, and and P.O. Brown, unpubl.). We anticipate that DNA microarrays of the kind described here will be useful in additional applications for which conventional dot blots, high-density gridded arrays on porous membranes, or FISH are currently used. These potential applica-

tions include comparative genomic hybridization (Kallioniemi et al. 1992), sequencing by hybridization (Drmanac et al. 1993), physical mapping of cloned or amplified sequences (Billings et al. 1991), and economical distribution of reagents for integrated genetic and physical mapping based on a common set of arrayed clones (Zehetner and Lehrach 1994).

METHODS

Amplification of Target DNA Elements

The array elements were prepared from physically mapped λ clones (Riles et al. 1993). The λ clones were amplified using randomly primed polymerase chain reaction (PCR) based on published and unpublished protocols (Bohlander et al. 1992; S. Nelson, unpubl.). The phage lysates were amplified in a 10-μl PCR reaction using 5 μM final concentration of primer A (GCTATCTTCAAGATCANNNNNN), 200 μM dNTPs, and 1 unit of *Taq* polymerase. Round A consisted of five cycles at 94°C for 1 min, 25°C for 1.5 min, 25–72°C over 7 min, and 72°C for 3 min using *Taq* polymerase (BMB). For round B, the reaction volume was brought up to 100 μl for a final concentration of 2 μM of primer B (GCTATCTTCAAGATCA), 200 μM dNTPs, and 4 units of *Taq* polymerase. Round B consisted of 30 cycles of 94°C for 1 min, 56°C for 2 min, and 72°C for 3 min. The amplification was performed in 96-well plates using crude phage lysates as the templates, resulting in an amplification of both the 35-kb λ vector and the 5-kb to 15-kb yeast insert sequences as a distribution of PCR products between 250 bp and 1500 bp in length.

The PCR products were purified and transferred into TE (10 mM Tris, 1 mM EDTA at pH 8.0) buffer using Sephadex G50 gel filtration (Pharmacia) and evaporated to dryness at room temperature overnight. Each of the 864 amplified λ clones was rehydrated in 15 μl of 3× SSC (20× SSC = 3 M NaCl, 0.3 M Na_3 citrate) in preparation for spotting onto the glass under normal room temperature conditions.

Preparation of DNA Microarrays

The microarrays were fabricated on poly-L-lysine coated microscope slides (Sigma). A custom-built arraying machine, consisting of four tweezer-like printing tips mounted 9 mm apart on a computer-controlled robotic stage (Shalon 1996), loaded 1 μl of the concentrated PCR product directly from corresponding clusters of four wells of 96-well storage plates and deposited ~5 nl of each sample onto each of 40 slides. Surface tension loaded the sample into the printing tip directly from the microwell plate and held the sample in the tip during the printing operation. Printing was achieved by lightly tapping the tip against the glass surface. The open-capillary design allowed for rapid rinsing and drying of the tips between samples. Figure 3 shows the layout of the arraying machine. Figure 4 shows a detailed view of the four printing tips and the staggered printing pattern on the microscope slides. Adjacent samples were spotted 380 μm apart on the slides. After each set of four samples was printed onto 40 slides, the printing tips were rinsed with a jet of water for 2 sec and then dried by lowering the tips onto a sponge for 2 sec. The process was repeated for all 864 samples and eight control spots.

After the spotting operation was complete, the slides were rehydrated in a humid chamber at room temperature for 2 hr, baked in an 80°C vacuum oven for 2 hr, then rinsed in 0.1% sodium dodecyl sulfate (SDS) to remove unadsorbed DNA. To reduce nonspecific adsorption of the labeled hybridization probe to the poly-L-lysine coated glass surface, the slides were treated with succinic anhydride. One gram of succinic anhydride was dissolved in 100 ml of 1-methyl-2-pyrrolidinone and then 100 ml of 0.2 M boric acid (pH 8.0) was added. The arrays were soaked in this solution for 10 min and then rinsed in distilled water four times for 5 min each. Immediately before use, the arrayed DNA elements were denatured by placing the slide in distilled water at 90°C for 2 min.

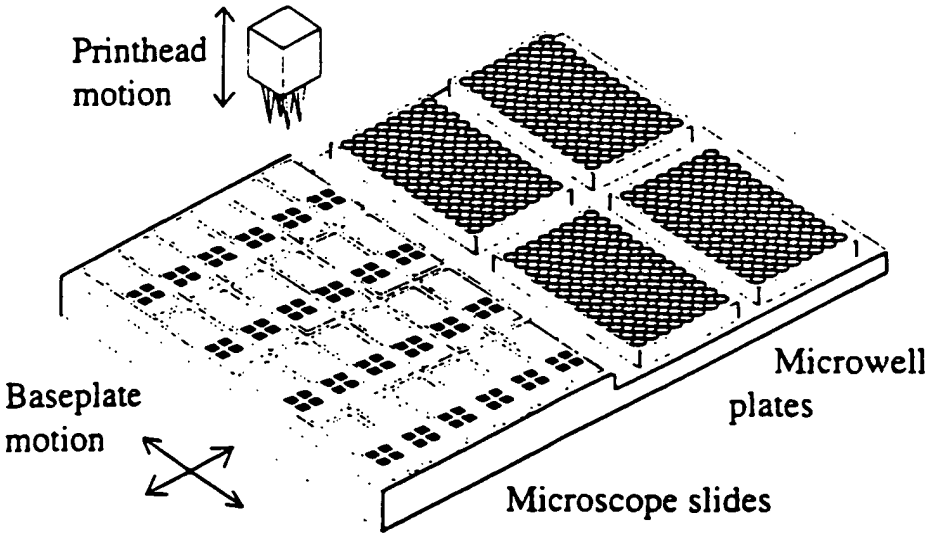


Figure 3 The layout of the arraying machine. All motions are under computer control. For more details of the arraying machine, see web page http://cmgm.stanford.edu/pbrown.

Amplification and Labeling of Hybridization Probe

The 16 chromosomes of *Saccharomyces cerevisiae* were separated using a contour-clamped homogeneous electric field (CHEF) agarose gel apparatus (Bio-Rad) (Chu et al. 1986). The 6 largest chromosomes were isolated in one gel slice and the smallest ten chromosomes in a second gel slice. The DNA from each slice was recovered using a gel extraction kit

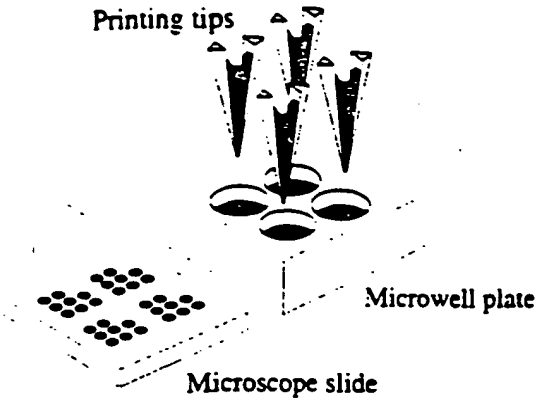


Figure 4 A close-up view of the four open-capillary printing tips. The tips are 9 mm apart and fit into four adjacent wells of a standard microwell plate and print arrays in a staggered fashion on microscope slides. For more details of the printing tips, see web page http://cmgm.stanford.edu/pbrown.

(Qiagen) and randomly amplified in a manner similar to that used in amplifying the target λ clones (Grothues et al. 1993). The main difference between this amplification procedure and the one used for the λ array elements is a filtration step between rounds A and B to remove primer-dimers and the use of a random 9-mer 3′ end on primer A. Following amplification, 2.5 μg of each of the amplified chromosome pools were separately random-primer labeled using Klenow polymerase (Amersham) with a lissamine-conjugated nucleotide analog (DuPont NEN) for the pool containing the 6 largest chromosomes and with a fluorescein-conjugated nucleotide analog (BMB) for the pool containing the smallest 10 chromosomes. The two fluorescent-labeled pools were mixed and concentrated using an ultrafiltration device (Amicon).

Hybridization

Five micrograms of the hybridization probe, consisting of both chromosome pools in 7.5 μl of TE, was denatured in a boiling water bath and then snap-cooled on ice. Concentrated hybridization solution (2.5 μl) was added to a final concentration of 5× SSC/0.1% SDS. The entire 10 μl of probe solution was transferred to the array surface, covered with a coverslip, placed in a custom-built single-slide humidity chamber, and incubated in a 60°C water bath for 12 hr. The custom-built waterproof slide chamber has a cavity just slightly bigger than a microscope slide and was kept at 100% humidity internally by the addition of 2 μl of water in a corner of the chamber. The slide was rinsed in 5× SSC/0.1% SDS for 5 min and then in 0.2× SSC/0.1% SDS for 5 min. All rinses were at room temperature. The array was then air dried, and a drop of antifade (Molecular Probes) was applied to the array under a 24-mm × 30-mm coverslip in preparation for scanning.

Detection and Analysis

A custom-built laser scanner was used to detect the two-color fluorescence hybridization signals from 1.8-cm × 1.8-cm arrays at 20-μm resolution. The glass substrate slide was mounted on a computer-controlled, two-axis translation stage (PM-500, Newport, Irvine, CA) that scanned the array over an upward-facing microscope objective (20×, 0.75NA Fluor, Nikon, Melville, NY) in a bidirectional raster pattern. A water-cooled Argon/Krypton laser (Innova 70 Spectrum, Coherent, Palo Alto, CA), operated in multiline mode, allowed for simultaneous specimen illumination at 488.0 nm and 568.2 nm. These two lines were isolated by a 488/568 dual-band excitation filter (Chroma Technology, Brattleboro, VT). An epifluorescence configuration with a dual-band 488/568 primary beam splitter (Chroma) excited both fluorophores simultaneously and directed fluorescence emissions toward the two-channel detector. Emissions were split by a secondary dichroic mirror with a 565 transition wavelength onto two multialkali cathode photomultiplier tubes (PMT; R928, Hamamatsu, Bridgewater, NJ), one with an HQ535/50 bandpass barrier filter and the other with a D630/60 bandpass barrier filter (Chroma). Preamplified PMT signals were read into a personal computer using a 12-bit analog-to-digital conversion board (RTI-834, Analog Devices, Norwood, MA), displayed in a graphics window, and stored to disk for further rendering and analysis. The back aperture of the 20× objective was deliberately underfilled by the illuminating laser beam to produce a large-diameter illuminating spot at the specimen (5-μm to 10-μm half-width). Stage scanning velocity was 100 mm/sec, and PMT signals were digitized at 100 μsec intervals. Two successive readings were summed for each pixel, such that pixel spacing in the final image was 20 μm. Beam power at the specimen was ~5 mW for each of the two lines.

The scanned image was despeckled using a graphics program (Hijaak Graphics Suite) and then analyzed using a custom image gridding program that created a spreadsheet of the average red and green hybridization intensities for each spot. The red and green hybridization intensities were corrected for optical cross talk between the fluorescein and lissamine channels, using experimentally determined coefficients.

ACKNOWLEDGMENTS

This research was supported by grant HG00450 from the National Institutes of Health–National Center for Human Genome Research, a National Science Foundation graduate fellowship to D.S., and by the Howard Hughes Medical Institute. P.O.B. is an assistant investigator of the Howard Hughes Medical Institute. We thank John Mulligan and John McCusker for help in preparing and amplifying the λ clones used in the arrays, Ren Xin Xia for writing the scanner control software and the image gridding and automatic karyotyping programs, Jeff van Ness at Darwin Molecular Corporation for suggesting the use of succinic anhydride, Stan Nelson, Linda McAllister, Joe deRisi, and Lolita Penland for helpful suggestions in the course of this work, and Joe deRisi and Linda McAllister for helpful comments on the manuscript.

The publication costs of this article were defrayed in part by payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 USC section 1734 solely to indicate this fact.

REFERENCES

Billings, P.R., C.L. Smith, and C.R. Cantor. 1991. New techniques for physical mapping of the human genome. *FASEB J.* **5:** 28–34.

Bohlander, S.K., R. Espinosa III, M.M. LeBeau, J.D. R wley, and M.O. Diaz. 1992. A method for the rapid sequence-independent amplification of microdissected chromosomal material. *Genomics* **13:** 1322–1324.

Chu, G., D. Vollrath, and R. Davis. 1986. Separation of large DNA molecules by contour clamped homogeneous electric fields. *Science* **234:** 1582–1585.

Drmanac, R., S. Drmanac, Z. Strezoska, T. Paunesku, I. Labat, M. Zeremski, J. Snoddy, W.K. Funkhouser, B. Koop, L. Hood, et al. 1993. DNA sequence determination by hybridization: A strategy for efficient large-scale sequencing. *Science* **260:** 1649–1652.

Grothues, D., C.R. Cantor, and C.L. Smith. 1993. PCR amplification of megabase DNA with tagged random primers (T-PCR). *Nucleic Acids Res.* **21:** 1321–1322.

Hudson, T.J., L.D. Stein, S.S. Gerety, J. Ma, A.B. Castle, J. Silva, D.K. Slonim, R. Baptista, L. Kruglyak, S.H. Xu, et al. 1995. An STS-based map of the human genome. *Science* **270:** 1945–1954.

Kallioniemi, A., O.P. Kallioniemi, D. Sudar, D. Rutovitz, J.W. Gray, F. Waldman, and D. Pinkel. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. *Science* **258:** 818–821.

Maskos, U. and E.M. Southern. 1992. Parallel analysis of oligodeoxyribonucleotide (oligonucleotide) interactions. I. Analysis of factors influencing oligonucleotide duplex formation. *Nucleic Acids Res.* **20:** 1675–1678.

Nelson, S.F., J.H. McCusker, M. Sander, Y. Kee, P. Modrich, and P.O. Brown. 1995. Genomic mismatch scanning: A new approach to genetic linkage mapping. *Nature Genet.* **4:**11–17.

Pease, A.C., D. Solas, E.J. Sullivan, M.T. Cronin, C.P. Holmes, and S.P. Fodor. 1994. Light-generated oligonucleotide arrays for rapid DNA sequence analysis. *Proc. Natl. Acad. Sci.* **91:** 5022–5026.

Riles, L., J.E. Dutchik, A. Baktha, B.K. McCauley, E.C. Thayer, M.P. Leckie, V.V. Braden, J.E. Depke, and M.V. Olson. 1993. Physical maps of the six smallest chromosomes of *Saccharomyces cerevisiae* at a resolution of 2.6 kilobase pairs. *Genetics* **134:** 81–150.

Ross, M.T., J.D. H heisel, A.P. Monac , Z. Larin, G. Zehetner, and H. Lehrach. 1992. High density gridded YAC filters: Their potential as genome mapping tools. In *Techniques for the analysis of complex genomes* (ed. Rakesh Anand), pp. 137–153. Academic Press, London, UK.

Schena, M., D. Shal n, R.W. Davis, and P.O. Brown. 1995. Quantitative monit ring of gene expression patterns with a complementary DNA microarray. *Science* **270:** 467–470.

Shalon, D., 1995. "DNA micro arrays: A new tool for genetic analysis." Ph.D. thesis, Stanford University, Stanford, CA.

Zehetner, G. and H. Lehrach. 1994. The reference library system—Sharing biological material and experimental data. *Nature* **367:** 489–491.

Received March 4, 1996; accepted in revised form May 9, 1996.